VOLUME 39 OCTOBER 1961 NUMBER 10

Canadian Journal of Physics

Published by THE NATIONAL RESEARCH COUNCIL

OTTAWA CANADA

CANADIAN JOURNAL OF PHYSICS

Under the authority of the Chairman of the Committee of the Privy Council on Scientific and Industrial Research, the National Research Council issues THE CANADIAN JOURNAL OF PHYSICS and five other journals devoted to the publication, in English or French, of the results of original scientific research. Matters of general policy concerning these journals are the responsibility of a joint Editorial Board consisting of: members representing the National Research Council of Canada; the Editors of the Journals; and members representing the Royal Society of Canada and four other scientific societies.

Manuscripts for publication should be submitted to Dr. H. E. Duckworth, Editor, Canadian Journal of Physics, Hamilton College, McMaster University, Hamilton, Ontario.
For instructions on preparation of copy, see **NOTES TO CONTRIBUTORS** (back cover).

Proof, correspondence concerning proof, and orders for reprints should be sent to the Manager, Editorial Office (Research Journals), Division of Administration and Awards, National Research Council, Ottawa 2, Canada.

Subscriptions, renewals, requests for single or back numbers, and all remittances should be sent to Division of Administration and Awards, National Research Council, Ottawa 2, Canada. Remittances should be made payable to the Receiver General of Canada, credit National Research Council.

The journals published, frequency of publication, and subscription prices are:

Canadian Journal of Biochemistry and Physiology	Monthly	$9.00 a year
Canadian Journal of Botany	Bimonthly	$6.00 a year
Canadian Journal of Chemistry	Monthly	$12.00 a year
Canadian Journal of Microbiology	Bimonthly	$6.00 a year
Canadian Journal of Physics	Monthly	$9.00 a year
Canadian Journal of Zoology	Bimonthly	$5.00 a year

The price of regular single numbers of all journals is $2.00.

Canadian Journal of Physics

Issued by THE NATIONAL RESEARCH COUNCIL OF CANADA

VOLUME 39 OCTOBER 1961 NUMBER 10

THE $T(\alpha,\gamma)Li^7$ REACTION[1]

G. M. GRIFFITHS, R. A. MORROW,[2] P. J. RILEY,[3] AND J. B. WARREN

ABSTRACT

The $T(\alpha,\gamma)Li^7$ reaction has been observed using thin targets of tritium absorbed in zirconium bombarded by singly charged helium ions of energy from 0.5 Mev to 1.9 Mev. The cross section rises smoothly with energy in a fashion characteristic of a direct radiative capture process. The ratio of the intensities of transitions to the first excited state at 478 kev and to the ground state is 0.40 ± 0.05. The angular distribution is, to within errors, isotropic at $E_\alpha = 0.8$ Mev but is significantly higher at 0° than at 90° for $E_\alpha = 1.6$ Mev. The total capture cross section at a mean alpha-particle energy in the target of 1.32 Mev is 3.58 ± 0.60 microbarns and the corresponding astrophysical S factor is 0.064 ± 0.016 kev barn in the center of mass system. The results are compared with recent theoretical results on the direct radiative capture process.

INTRODUCTION

In the last few years there has been considerable interest in direct reactions which take place without the formation of a compound nucleus and which, in many cases, depend on the nuclear wave function at and even beyond the conventional nuclear radius. A particular class of these reactions is the direct radiative capture process. This process contributes a significant gamma-ray yield in the energy regions between resonances for a number of reactions at low bombarding energy, such as $O^{16}(p,\gamma)F^{17}$ (Warren *et al.* 1954), $Ne^{20}(p,\gamma)Na^{21}$ (Tanner 1959), and $T(\alpha,\gamma)Li^7$ (Riley *et al.* 1958 and Griffiths 1959).

We have extended the earlier measurements made in this laboratory on the $T(\alpha,\gamma)Li^7$ reaction. The reaction has also been studied by Holmgren and Johnston (1959), who studied the mirror reaction $He^3(\alpha,\gamma)Be^7$ as well. The cross sections are over 2000 times greater than earlier estimates given by Cameron (1957) based on work of Salpeter (1952), which did not consider the direct reaction mechanism. For this reason these reactions, particularly the mirror reaction, are of importance to nucleogenesis since they provide a second method of bypassing the mass 5 nuclei as pointed out by Fowler (1958) and Cameron (1958). In addition, the results are of interest because of recent developments in direct capture theory by Christy and Duck (1961) and by Tombrello and Phillips (1961). The latter authors have interpreted the

[1]Manuscript received July 5, 1961.
Contribution from the Physics Department, University of British Columbia, Vancouver 8, B.C.
[2]Present address: Palmer Physical Laboratory, Princeton, N.J.
[3]Present address: Physics Department, University of Alberta, Edmonton, Alberta.

radiative capture in terms of a "two-body cluster" model for Li^7 and Be^7 which assumes that at low excitation the mass 7 nuclei can be thought of as an alpha particle plus a mass 3 particle in a relative P state.

EXPERIMENTAL

Targets of tritium absorbed in thin evaporated layers of zirconium on platinum backings, kindly supplied by the Oak Ridge National Laboratories, were bombarded with 2 to 8 microamperes of singly charged helium ions from the University of British Columbia Van de Graaff generator in the energy range from 0.5 to 2.0 Mev. The capture gamma rays were observed with well-shielded sodium iodide scintillation counters. Because of the low gamma-ray yield, backgrounds, particularly that from $C^{13}(\alpha,n)O^{16}$, were reduced by thorough cleaning of parts struck by the beam and by placing a liquid air trap in the beam tube just in front of the target holder.

Scintillation Counter Efficiency

The counter used in most of this work consisted of a $4\frac{1}{2}$ in. long by $2\frac{3}{4}$ in. diameter sodium iodide crystal mounted on a Dumont K1213 photomultiplier operated at 900 volts, with pulse height analysis provided by a 100 channel Computing Devices of Canada kicksorter. Five inches of lead shielding were used on the counter to reduce the background from the generator and from radiothorium and K^{40} in the concrete walls of the laboratory. Provision was made to rotate the counter assembly through $\pm 135°$ to the incident beam direction.

A theoretical and experimental study of the efficiency of this counter for γ-rays in the energy range from 0.5 Mev to 20 Mev has been made with the results shown in Fig. 1. Although the $T(\alpha,\gamma)Li^7$ gamma rays lie in the energy

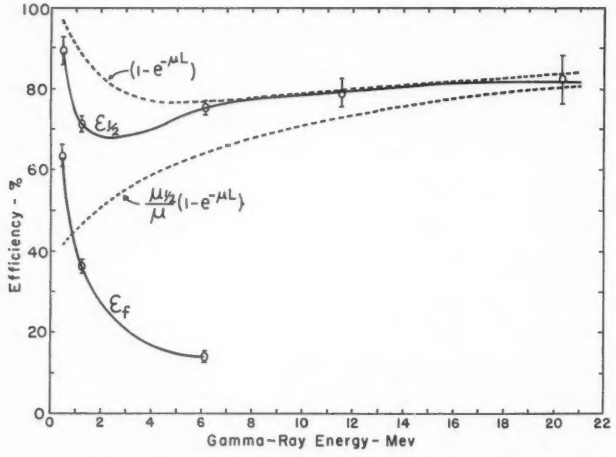


FIG. 1. Gamma-ray efficiency for the $2\frac{3}{4}$ in. diameter by $4\frac{1}{2}$ in. long NaI counter showing the efficiency for a bias equal to half the γ-ray energy, $\epsilon_{\frac{1}{2}}$, and for a bias set just below the full energy peak, ϵ_f. The dotted curves show the total theoretical primary absorption (upper curve) and the part of the total theoretical primary absorption which initially leaves more than half the γ-ray energy in the crystal.

region from 0.4 to 3.5 Mev the efficiency at 20 Mev was also required because the tritium content of the targets was measured in terms of the yield of 20-Mev gamma rays from the $T(p,\gamma)He^4$ reaction. The efficiency, ϵ_b, used here is defined for a given bias on the counting system by the following relation for the number of counts above the bias,

$$N_c = \left(\frac{dN_\gamma}{d\omega}\right)\frac{A\epsilon_b}{(d+e)^2}$$

where the first factor is the number of gamma rays per unit solid angle, A is the area of the end face of the counter, d is the distance from the end face to the gamma-ray source, and e is the "effective center" distance. This distance is measured from the end of the crystal to a point inside the crystal such that the counting rate varies as the inverse square of the distance of the gamma-ray source from this point. "e" is a function of gamma-ray energy and is obtained from the intercept on the distance axis of plots of the square root of the counting rate versus distance. These plots are remarkably straight lines for distances of several meters to less than the length of the crystal. Figure 1 shows the efficiency, $\epsilon_{\frac{1}{2}}$, for a bias equal to half the γ-ray energy as used in most of this work, as well as the efficiency for a bias set just below the full energy peak, ϵ_f. The two lower energy points were obtained in terms of absolutely calibrated sources of Na^{22} and Co^{60}. The 6.14-Mev point was determined by simultaneous counting of alpha particles from the $F^{19}(p,\alpha,\gamma)O^{16}$ reaction with a thin window proportional counter as suggested by Van Allen and Smith (1941). The upper part of the $\epsilon_{\frac{1}{2}}$ curve was obtained by intercomparison with a standard brass geiger counter similar to that used and calibrated by Barnes *et al.* (1952). Details of this work will be published at a later date.

Tritium Targets

The characteristics of the four targets used are given in Table I. These

TABLE I

Target tritium content

Number	Tritium content, 10^{18} atoms/cm^2			Ratio, T/Zr	ΔE (kev)	
	By zirconium	By β-rays	By γ-rays		0.8-Mev protons	1.5-Mev alphas
1	1.73	1.75	1.72	1.0	32.5	190
2	—	—	14.9	—	—	—
3	1.35	1.30	2.41	1.8	25.5	148
4	3.53	2.97	3.81	1.1	66	385

consisted of a weighed thickness of zirconium evaporated onto 0.01 in. thick platinum disks and saturated with tritium. They were fixed to a thin copper backing mounted on a cooled copper rod. There was no appreciable loss of yield after as much as 5 hours running with 5 microamperes of 1.5-Mev alphas on a $\frac{1}{4}$-in. spot. The uniformity of tritium surface density was checked by observing the $T(p,\gamma)He^4$ yield at various places on the surface and found to be better than 5%.

The tritium content was estimated by the suppliers in terms of the zirconium thickness assuming a one to one atomic ratio of T to Zr, and by measuring the ionization produced by the 18-kev beta activity from the target. The suppliers indicated that the atomic ratio could vary by more than 50% and for this reason we have made independent estimates of the tritium content as shown under the heading "by γ-rays" in Table I. First, the relative tritium contents were obtained by measuring the neutron yield from the $T(p,n)He^3$ reaction at 1.5 Mev using a γ-ray insensitive plastic phosphor neutron counter (Griffiths *et al.* 1959). Second, the tritium was estimated absolutely in terms of the 20-Mev γ-ray yield from the $T(p,\gamma)He^4$ reaction at 800 kev, this energy being chosen well below the 873-kev resonance in fluorine because the fluorine contamination of the targets was appreciable. The absolute tritium content was based on the $T(p,\gamma)He^4$ cross section given by Perry and Bame (1955) for the mean proton energy in the target (2.6×10^{-30} cm^2/steradian at 90° for target number 4 at 45° to the beam and a mean proton energy of 755 kev). The final two columns in Table I give the energy loss for protons and alpha particles passing perpendicularly through the target, assuming a one to one atomic ratio and using figures interpolated from the stopping cross-section data given by Whaling (1958).

A typical 20-Mev γ-ray spectrum is shown in Fig. 2. Since the $T(p,\gamma)He^4$

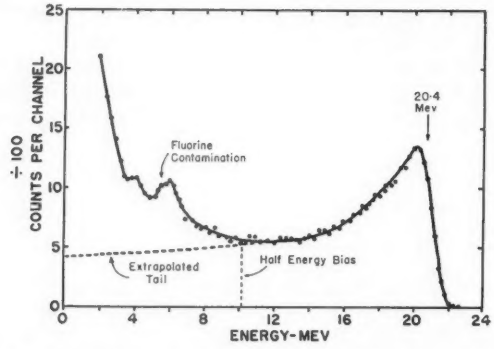


FIG. 2. Typical $T(p,\gamma)He^4$ spectrum used for estimating the tritium content of the targets showing the half energy bias and extrapolated tail.

cross section obtained by Perry and Bame (1955) using a gas target was based on a similar scintillation counter spectrum obtained with a crystal of comparable size we have extrapolated the tail of the spectrum to zero pulse height as shown and assumed that the number of counts under this spectrum corresponded to the total theoretical primary absorption in the crystal, as was done by Perry and Bame. For these measurements the counter was 35 centimeters from the target, therefore only small solid angle corrections were required. The gamma-ray flux estimated by this method is 25% higher than that estimated in terms of the efficiency for half energy bias shown in Fig. 1 for 20-Mev γ-rays obtained by intercomparison with the standard geiger counter.

This suggests that the tail of the spectrum as shown contains considerably more pulses than can be accounted for by primary absorption events in the crystal. This discrepancy, although it may introduce some uncertainty in the $T(p,\gamma)He^4$ cross section, should be cancelled in deducing the tritium content of the target since we have made a similar interpretation of the scintillation counter spectrum for a comparable sized counter as was done in establishing the cross section. The ratio of tritium content estimated by γ-rays to that estimated from the zirconium thickness, assuming a one to one atomic ratio, is shown in Table I. The agreement is very good for targets 1 and 4 which were used for the absolute yield measurements. Target 3 was used for angular distribution and energy measurements and target 2, a thick target, was used for investigating secondary neutron effects discussed below. A decay correction of nearly 10% was made for the tritium decay between the time target 1 was made and when it was used; it is not known if the He^3 daughter of the decay escapes from the target.

RESULTS

Gamma-ray Spectra

Figure 3 shows typical γ-ray spectra observed for alpha-particle bombarding energies of 1.6 Mev and 0.8 Mev with beam currents of 4 to 5 microamperes

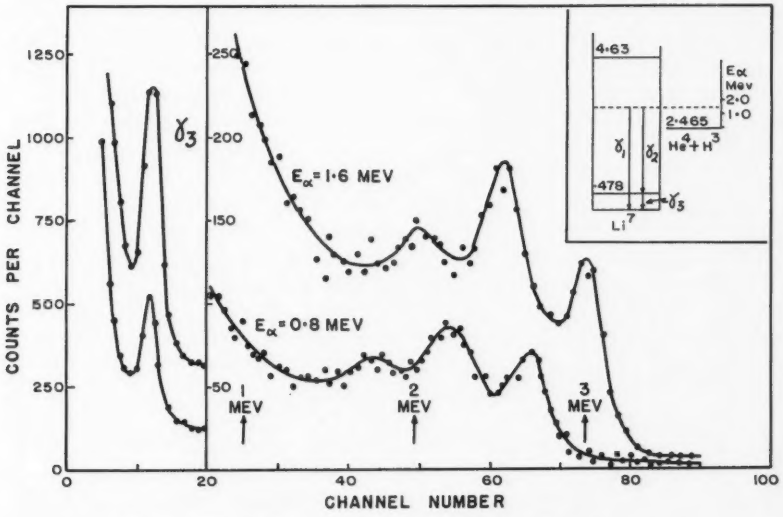


FIG. 3. $T(\alpha,\gamma)Li^7$ γ-ray spectra at two bombarding energies.

on target number 4. The inset shows the energy region of interest. In the energy region from 2 to 3 Mev a background of approximately 24 counts per channel has been subtracted from the 1.6-Mev curve, which was obtained in 20 minutes with an integrated current of 5000 microcoulombs. In this

1-Mev energy region the time-dependent background was 8 counts per minute and the beam-dependent background was 22 counts per minute at $E_\alpha = 1.6$ Mev and 9 counts per minute at $E_\alpha = 0.8$ Mev as measured on a clean piece of zirconium. The mean alpha-particle energies in the target for these runs were $\bar{E}_\alpha = 1.32$ Mev and $\bar{E}_\alpha = 0.56$ Mev with the target at 45° to the incident beam.

In addition to the three $T(\alpha,\gamma)Li^7$ γ-rays it is apparent that there is a substantial background from the target particularly at low energies. Although the presence of two overlapping γ-ray spectra of 3.0 and 2.5 Mev may not be clear in Fig. 3 because the separation in energy of the two γ-rays, γ_1 and γ_2, is almost the same as the separation of the multiple peaks produced by each γ-ray, the presence of the 2.5 Mev, γ_2, to the first excited state in Li^7 is confirmed by the presence of the 0.477-Mev ground state transition, γ_3, from the first excited state. In addition, the spectra do not have the shape characteristic of a single γ-ray in this energy region. In order to separate the contributions to the spectra from γ_1 and γ_2, standard spectra were obtained from radiothorium (2.62 Mev), $C^{12}(d,p,\gamma)C^{13}$ (3.09 Mev), $B^{11}(p,\gamma)C^{12}$ (4.43 Mev), and $F^{19}(p,\alpha,\gamma)O^{16}$ (6.14 Mev). From these, standard spectra could be constructed for intermediate energies. The peaks of the standard spectra had to be broadened slightly from the shapes obtained for monochromatic γ-rays in order that the decomposition procedure would result in a reasonably smooth background, because the $T(\alpha,\gamma)Li^7$ spectra are broadened due to the appreciable target thickness. The procedure was to fit a standard spectrum to the upper photopeak for γ_1, with allowance for the background above this peak and then to subtract it. Finally a γ_2 spectrum was subtracted, leaving a smooth background. A decomposition of this type is shown in Fig. 4 for data

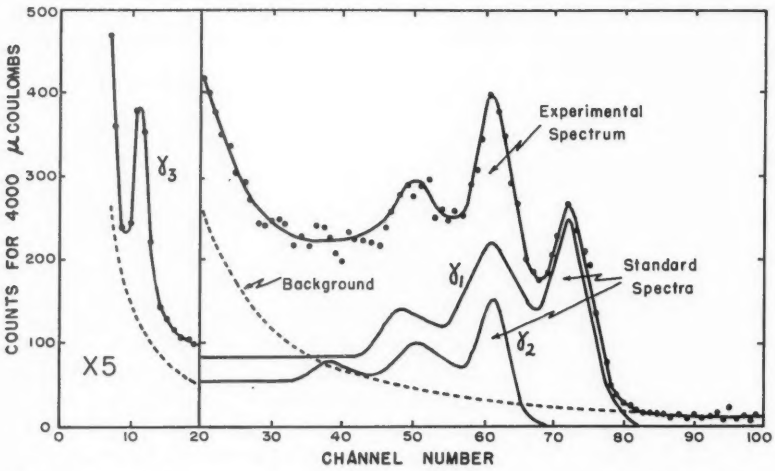


FIG. 4. $T(\alpha,\gamma)Li^7$ γ-ray spectrum at $E_\alpha = 1.6$ Mev showing the decomposition into components and the remaining background (broken line).

taken at a bombarding energy of 1.6 Mev showing the residual background which comes from the target, since ordinary time-dependent and beam-dependent backgrounds have been subtracted.

The γ-ray energy was found to vary as $(Q+\frac{3}{7}E_\alpha)$ where E_α is the mean alpha energy in the target. The mean alpha energies calculated from the observed γ-ray energies agreed to better than 2% with those calculated from the target thickness data. A rough estimate of the intensity of γ_3 agreed with the intensity of γ_2 obtained after separating the spectra, confirming that these two γ-rays are in cascade.

Measurements of the γ-ray energy as a function of angle were made at 0°, 45°, 90°, and 135° for a bombarding energy of 1.6 Mev. These clearly showed the expected Doppler shift amounting to 42 ± 5 kev for γ_1 between 90° and 0° in good agreement with a calculated shift of 45 kev.

Branching Ratio

From the decomposed spectra and a knowledge of the energy dependence of the counter efficiency it was possible to determine the ratio of intensities of the transitions to the first excited state and to the ground state. Within the experimental accuracy of about 5% the ratio was independent of angle indicating that γ_1 and γ_2 have practically the same angular distribution. The branching ratios are

$$\gamma_2/\gamma_1 = 0.45\pm0.03 \text{ at } 1.6 \text{ Mev } (\bar{E}_\alpha = 1.32 \text{ Mev}),$$

$$\gamma_2/\gamma_1 = 0.36\pm0.03 \text{ at } 0.8 \text{ Mev } (\bar{E}_\alpha = 0.56 \text{ Mev}).$$

Angular Distributions

On the assumption of a direct radiative process one would expect the $T(\alpha,\gamma)Li^7$ reaction at low energies to proceed by electric dipole radiation from a continuum S state (S wave α-particles) to the $j = \frac{3}{2}-$ ground state and the $j = \frac{1}{2}-$ first excited state. In this case the angular distributions should be isotropic. The preliminary results of Riley *et al.* (1958) indicated that at 1.6-Mev bombarding energy the yield at 0° was greater than at 90°. For this reason more detailed angular distribution results have been obtained at 1.6 Mev and 0.8 Mev. Only the ground state radiation γ_1 was analyzed since the branching ratio is independent of angle to within the accuracy of about 5%.

At 800 kev ($\bar{E}_\alpha = 0.56$ Mev), measurements were made at 0° and 90° only, giving a 0° to 90° yield ratio 1.05 ± 0.06. Thus at this energy the angular distribution is isotropic within the errors of the measurements.

At 1.6 Mev, spectra were obtained for γ_1 at 0° and 45°, 90° and 135° on each side of the target. The target plane was positioned either at 90° or at ±45° to the beam direction, depending on the angle of observation, and corrections were made for target thickness and absorption in the backing due to the change in target angle. The results for six different runs at each angle were averaged to give the intensity ratios shown in Table II. The errors given are based on counting statistics only but they are close to the root mean square errors of the individual runs from the mean for each angle. The results were symmetrical on either side of 0° to better than 4%. Finally small

TABLE II

Angular distribution ratios for γ_1 at 1.6 Mev

Intensity ratio	Experimental	Corrected
0°/90°	1.30±0.04	1.28±0.04
45°/90°	1.08±0.06	1.06±0.06
135°/90°	1.04±0.06	1.02±0.06

corrections were applied for the finite solid angle of the detector and for center of mass motion to give the corrected ratios shown in Table II. It should be pointed out that the target-dependent background left over after separating the standard spectra for γ_1 and γ_2 was not itself isotropic; this is discussed in the next section. However, it is felt that the decomposition procedure used for the spectra has effectively removed any effect of the background anisotropy from the angular distribution for γ_1.

Neutron Background

The target-dependent background, which was subtracted from the data taken for the 1.6-Mev angular distributions, was very markedly anisotropic with a distribution that could be fitted by the function $(1+0.35 \cos \theta)$. In addition, the spectrum of this background which remained after separating out the standard γ-ray spectra, was much like the background produced in the sodium iodide counter by neutrons. It was unlikely that the $C^{13}(\alpha,n)O^{16}$ reaction contributed appreciably as independent checks on the $C^{12}(p,\gamma)N^{13}$ and $C^{13}(p,\gamma)N^{14}$ reactions indicated a very low carbon contamination in or on the target. Using a thick target (number 2) bombarded by 1.6-Mev α-particles, the neutron spectrum, yield, and angular distribution were investigated with a γ-ray insensitive neutron counter (Griffiths *et al.* 1959). This counter has poor energy resolution, but by comparison with monoenergetic calibration spectra and Ra–Be neutron spectra it was concluded that the neutrons from the tritium target had energies between 5 and 10 Mev. This seems to rule out the possibility that the neutrons were produced by a small amount of molecular deuterium in the mass 4 beam which would have produced 14-Mev neutrons by the $T(d,n)He^3$ reaction. It seems most likely that the neutrons observed are a result of a secondary process; tritons scattered by incident α-particles collide with further target tritium atoms to produce $T(t,2n)He^4$ neutrons similar to the secondary neutron production which has been observed in heavy ice targets under proton bombardment (Singh *et al.* 1959). The anisotropy of the neutron yield from target number 4 at 1.6 Mev was measured and found to agree to within 10% with the angular distribution of the target-dependent background observed in the scintillation counter. In addition, a rough check showed that the ratio of the neutron production at 1.6 Mev to that at 0.8 Mev was also in agreement with the ratio of the residual scintillation counter backgrounds at these energies. These results then support the view that the greatest part of the target-dependent background is due to neutrons from the target, likely produced by scattered tritons reacting with other target tritons. This secondary neutron effect should be even greater for a gas target where there are no additional zirconium atoms to contribute to stopping the recoil tritons.

Cross Section

The absolute 90° differential cross section for the T(α,γ)Li⁷ reaction at 1.6 Mev was obtained with targets 1 and 4 at 45° to the incident α-particle beam. The runs were interspaced with T(p,γ)He⁴ runs to establish the tritium content of the targets at the same time and on the same target position. The counter solid angle was 0.18 steradian measured from the target to the "effective center" of the scintillation counter, 4.8 centimeters inside the front face of the counter. This effective center distance was determined roughly by calculating the distance inside the crystal at which half the total absorption occurred and more accurately by finding the intercept on the distance axis for a plot of the reciprocal of the square root of the counting rate versus distance from the front end of the crystal for a number of γ-ray sources. The use of an "effective center" permits one to define an efficiency factor independent of the distance of the source to an accuracy better than 5% for distances of the source about equal to the length of the crystal. The efficiency of the counter was approximately 0.68±0.03 for γ-rays in the energy range from 2.5 to 3 Mev, where this efficiency is the ratio of the number of counts in the spectrum above a bias equal to half the γ-ray energy to the number of γ-rays entering the counter solid angle. In general, runs at energies other than 1.6 Mev were taken just prior to or just after runs at 1.6 Mev. Runs taken with targets 1 and 4 agreed to within 10%. Several runs taken on target 1 with a sodium iodide counter 3½ in. long by 2½ in. diameter agreed to within 15% with yield data from the larger counter. The results are collected in Table III and in Figs. 5 and 6. The first column in the table lists the mean alpha-particle energy in the target, the next three columns the 90° differential

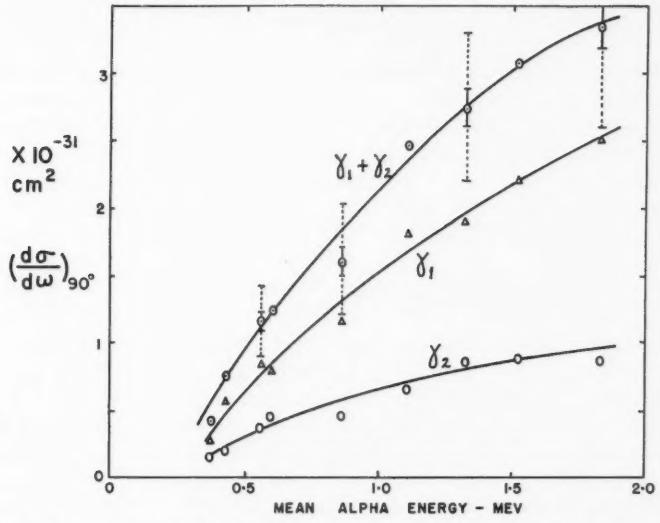


FIG. 5. The T(α,γ)Li⁷ differential cross section measured at 90° to the incident beam. The solid vertical lines indicate the statistical uncertainties while the dotted lines indicate the estimated total standard errors.

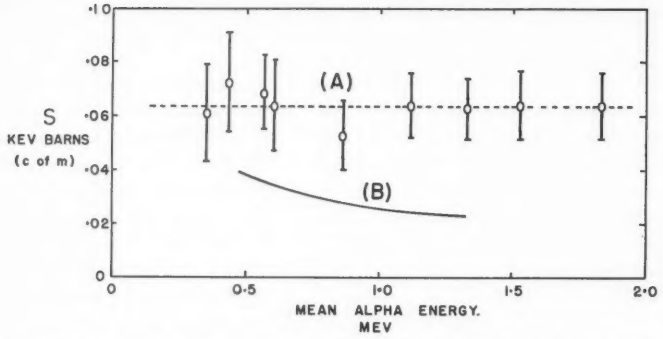


FIG. 6. The center of mass cross section factor S for the $T(\alpha,\gamma)Li^7$ reaction. Curve A shows the present results and curve B shows approximately the S value obtained by drawing a smooth line through the points given by Holmgren and Johnston (1959).

TABLE III

$T(\alpha,\gamma)Li^7$ cross-section results

	$(d\sigma/d\omega)_{90°}\times10^{31}$ cm²/sterad				
$\bar{E}$	γ_1	γ_2	$\gamma_1+\gamma_2$	σ_{total} $\times10^{30}$ cm²	S
1.83	2.50	.85	3.35±0.70	4.40±0.9	.064
1.52	2.20	.85	3.05±0.60	4.00±0.8	.0645
1.32	1.89	.85	2.74±0.50	3.58±0.6	.0632
1.11	1.80	.65	2.45±0.50	3.15±0.6	.0645
.86	1.14	.45	1.59±0.40	2.00±0.5	.0530
.61	.78	.45	1.23±0.30	1.55±0.4	.0645
.56	.84	.31	1.15±0.30	1.43±0.3	.0690
.43	.56	.19	0.75±0.20	.95±0.25	.0730
.35	.26	.14	0.40±0.13	.50±0.15	.0615

cross sections, and the fourth column lists the total cross section integrated over all angles taking into account a small correction for the nonisotropic yield at higher energies. The final column lists the center of mass value of the cross-section factor S used in astrophysical work (Burbidge *et al.* 1957)

$$S = \sigma(E_1)E_1\frac{A_0}{A_0+A_1}\exp(31.28\,Z_1Z_2A_1^{\frac{1}{2}}E_1^{-\frac{1}{2}})$$

$$\approx \sigma(\bar{E})\frac{3}{7}\bar{E}\exp(125.12\,\bar{E}^{-\frac{1}{2}})\text{ kev barns},$$

where $\bar{E}$ is the mean alpha-particle energy in the target in kev. The value of S as a function of energy is shown in Fig. 6. Within the accuracy S is a constant with a mean value of 0.064 kev barns and a mean square deviation of 8% from the mean. The estimated errors, however, are considerably greater than this and amount to 20 to 25% as shown in Table III and Fig. 5. These estimated errors are in all cases greater than the statistical error of about 5% for points above 800 kev and about 8% for points below 800 kev. The main contributions to the estimated error are 15% for target tritium content, 5% for statistics,

5% for scintillation counter efficiency, and 5% for counter solid angle. The beam current measurements were made with a current integrator to an accuracy of 2%, and the targets were run at +100 volts to prevent secondary electron emission.

DISCUSSION

The total cross section obtained from the present results is a factor of 2 to 2.5 higher than that obtained by Holmgren and Johnston (1959) over most of the energy range measured, and there is not as much evidence for an increase in the S factor as the bombarding energy is decreased. It should be pointed out that small errors in estimating the mean energy of the alpha particles in the target cause increasingly large errors in the S value as the energy is lowered because of the strong energy dependence of the penetrability. Using our ratio γ_2/γ_1 and our total $S = 0.064$ kev barns, the S factor for the transition to the ground state of Li^7 is $S_{3/2} = 0.046$ kev barns in surprisingly good agreement with the value calculated by Christy and Duck (1961), $S_{3/2} = 0.05$ kev barns.

The observed ratio of the number of gamma rays to the first excited state to the number to the ground state is 0.45 at 1.6 Mev and 0.36 at 0.8 Mev in good agreement with the value 0.42 obtained by Tombrello and Phillips (1961) using direct capture theory. In addition, the theory given by these authors combined with the present cross section leads to a reduced width $\theta^2_{3+4} = 0.12$ for the configuration alpha particle plus triton and, as pointed out by these authors, this is much larger than the reduced width θ^2_{1+6}, which supports the assumption of a two-particle cluster structure for Li^7. No detailed analysis of the angular distribution data, particularly at higher bombarding energies where there are pronounced departures from isotropy, have been undertaken yet. The main process consisting of the capture of S-wave alpha particles to form a relative p-state of the alpha and triton with electric dipole emission leads to an isotropic distribution characteristic of the results at low bombarding energies. Tombrello and Phillips have calculated the contribution due to magnetic dipole radiation following p-wave capture and find the cross section at $E_\alpha = 2$ Mev to be 0.5×10^{-31} cm^2. This is only just over 1% of the total cross section given in Table III for $E_\alpha = 1.8$ Mev. More detailed calculations are now being carried out to check further into this matter.

REFERENCES

Barnes, C. A., Carver, J. H., Stafford, G. H., and Wilkinson, D. H. 1952. Phys. Rev. **86**, 359.

Burbidge, E. M., Burbidge, G. R., Fowler, W. A., and Hoyle, F. 1957. Revs. Modern Phys. **29**, 547.

Cameron, A. G. W. 1957. Report No. AECL-454.

——— 1958. Bull. Am. Phys. Soc. II, **3**, 227.

Christy, R. F. and Duck, I. 1961. Nuclear Phys. **24**, 89.

Fowler, W. A. 1958. Astrophys. J. **127**, 551.

Griffiths, G. M. 1959. Compt. rend. congr. intern. phys. nucléaire, Dunod, Paris, 1959, page 447.

Griffiths, G. M., Singh, P. P., Ssu, Y. I., and Warren, J. B. 1959. Can. J. Phys. **37**, 858.

Grodstein, G. W. 1957. Natl. Bur. Standards Circ. 583, April 30, 1957.

Holmgren, H. D. and Johnston, R. L. 1959. Phys. Rev. **113**, 1556.

PERRY, J. E. and BAME, S. J. 1955. Phys. Rev. **99**, 1368.
RILEY, P., WARREN, J., and GRIFFITHS, G. 1958. Bull. Am. Phys. Soc. II, **3**, No. 5, 330.
SALPETER, E. E. 1952. Phys. Rev. **88**, 547.
SINGH, P. P., GRIFFITHS, G. M., SSU, Y. I., and WARREN, J. B. 1959. Can. J. Phys. **37**, 866.
TANNER, N. 1959. Phys. Rev. **114**, 1060.
TOMBRELLO, T. A. and PHILLIPS, G. C. 1961. Phys. Rev. **122**, 224.
VAN ALLEN, J. and SMITH, N. 1941. Phys. Rev. **59**, 618.
WARREN, J. B., LAURIE, K. A., JAMES, D. B., and ERDMAN, K. L. 1954. Can. J. Phys. **32**, 563.
WHALING, W. 1958. Handbuch der Physik, **34**, 193.

THE EFFECTS OF ENERGETIC TRAPPED PARTICLES ON MAGNETOSPHERIC MOTIONS AND IONOSPHERIC CURRENTS[1]

J. A. FEJER[2]

ABSTRACT

It is shown that the temporary distortion of a previously stable belt of energetic charged particles, trapped in the earth's magnetic field, results in a system of ionospheric currents and magnetospheric motions. The currents are caused by the charge separation which occurs in the course of the subsequent adiabatic motion of the energetic particles. The resulting current system closely resembles the well-known *Ds* current system. The polarization fields, which drive these currents, give rise to magnetospheric motions similar to those of Axford and Hines and the implications of their work then also result. The distortion of the belt could be caused either directly, by the introduction of a new group of trapped particles mainly on the day side of the magnetosphere, or indirectly, through the relatively sudden compression of the earth's field on the day side, caused by the arrival of a neutral stream of charged particles from the sun.

1. INTRODUCTION

In most recent theories of magnetic storms the sequence of observed events is attributed to the arrival of a neutral stream of charged particles from the sun or to a great intensification of such a stream. The changes in the magnetic field, which take place during a magnetic storm, are often interpreted in terms of upper atmospheric current systems, although the location of these external current systems cannot be uniquely determined from measurements on the ground.

Thus the decrease of the field observed on the ground during the main phase of a magnetic storm has been interpreted by Schmidt (1924) and later by Chapman and Ferraro (1931) as a ring current in the equatorial plane at a distance of a few earth radii. Such a ring current was postulated earlier by Stormer (1911) to explain the position of the auroral zones with his theory.

Likewise the diurnal changes in the storm field, averaged over many magnetic storms, have been interpreted as a current system at about *E*-region height, the *Ds* current system (Chapman 1935). Similar current systems were obtained for magnetic bays by Silsbee and Vestine (1942).

Singer (1957) suggested that the equatorial drift (westward for ions and eastward for electrons) of energetic charged particles is the cause of the ring current and the main phase. The presence of energetic charged particles has been subsequently confirmed by observations with the aid of satellites (Van Allen and Frank 1959).

Dessler and Parker (1959) proposed that the trapping of such particles is made possible by instabilities at the boundary separating a solar stream from the plasma confined by the earth's magnetic field. The instabilities make the penetration and subsequent diffusion of neutral plasma blobs into the earth's field possible in a relatively short time.

[1]Manuscript received July 10, 1961.
Contribution from the D.R.B. Theoretical Studies Group, Ottawa, Canada.
[2]Present address: Defence Systems Division, General Motors Corporation, Santa Barbara, California.

It is generally assumed that the plasma surrounding the earth is confined by the earth's magnetic field up to distances of the order of 10 earth radii. Measurements with the aid of Explorer VI and Pioneer V (Smith *et al.* 1960) show that the earth's field does extend to distances of this order but that there are systematic differences between the values of the measured field and the field extrapolated from the ground. These differences were tentatively interpreted by Smith *et al.* in terms of a toroidal ring current of about 5 million amperes at a distance of about 8 earth radii from the center of the earth to the center of the current.

Gold (1959) has used the word 'magnetosphere' to describe the plasma surrounding the earth and dominated by the earth's magnetic field. He discussed the type of motions possible in the magnetosphere.

Axford and Hines (1961) have examined the nature of magnetospheric motions caused by an external neutral stream of charged particles. They conclude that a viscous-like transfer of momentum from the external stream causes a certain pattern of motions inside the magnetosphere, and that this pattern is significant to many geophysical phenomena. They conclude in particular that a current system similar to the *Ds* system follows from these magnetospheric motions.

In the present paper a mechanism is described which leads to ionospheric current systems similar to the *Ds* current system and to a pattern of magnetospheric motions similar to that of Axford and Hines (1961). The mechanism is based on the assumption that occasionally the axial symmetry of the distribution of energetic charged particles trapped in the magnetosphere is temporarily disturbed (for simplicity the earth's unperturbed magnetic field is regarded here as axially symmetrical) and that symmetry is re-established after some hours mainly by the same equatorial drifts of energetic charged particles (with a spread in velocities appropriate to the energy spectrum) which according to Singer (1957) causes the ring current and the main phase. Such periods of temporary asymmetry may well occur several times in succession during a magnetic storm. The ionospheric currents and magnetospheric motions occur only while the asymmetry persists, although a ring current at great distances continues to flow after the re-establishment of symmetry.

Two possible causes of temporary asymmetry are suggested in the present paper. The first of these is the preferential trapping of a group of energetic charged particles on the day side. The second possible cause is taken to be a relatively sudden distortion of the earth's magnetic field by a solar stream. Although this latter process destroys the initial symmetry of both the magnetic field and the particle distribution, its effects are shown to be similar to those of the first process in which the symmetry of the magnetic field is at least initially preserved.

The nature of possible motions in the magnetosphere is first discussed from a general point of view in Section 2 of this paper. The consequences of an initial asymmetry in the distribution of trapped energetic charged particles with respect to the magnetic field are examined in Sections 3 and 4 with the aid of the concepts developed in Section 2.

2. MOTIONS OF THE MAGNETOSPHERE

There seem to be two different types of ionized component gases present in the magnetosphere in addition to a certain amount of neutral gas. The first type of gas is of relatively low energy; its temperature is believed to be about 1000–2000° K. Its density is known approximately up to distances of about 6 earth radii. The density below the peak of the F_2 layer is obtained from ionosonde data. At heights between the F_2 layer peak and about 1000 km the density is obtained from satellite observations of different type and from observations of incoherent scattering. At still greater heights of up to 6 earth radii the observation of nose whistlers (Pope 1961) yields electron densities.

In addition to this low energy (thermal) component there are energetic trapped charged particles, some of which have been observed with the aid of satellite-borne counters (Van Allen and Frank 1959). Some of these particles are believed to be of solar origin; the consequences of preferential trapping of groups of such particles on the day side will be discussed in Section 3.

The motions of the low energy component in the absence of energetic charged particles will be considered first. It follows from the frozen field concept that only such motions of the low energy component of the magnetosphere are permitted in which ionized matter, which at one time occupies a tube of force, continues at all later times to occupy a tube of force of equal flux content. As Gold (1959) has pointed out, the earth's field must remain almost unaltered by the motions.

The motions of the low energy component can have a number of possible causes. One of these is the tidal wind system in the neutral air which is believed to cause the dynamo current system. Associated with the dynamo current system there is an electric field due to polarization charges and derivable from a potential. The global potential distribution over the thin spherical shell, in which the dynamo currents flow, is obtained from the dynamo theory (for example, Fejer 1953); it may be extended from this shell (the dynamo region) into the magnetosphere because the magnetic field lines are also lines of equipotential.* If a knowledge of the electric polarization field $\mathbf{E}$ is assumed then the drift velocity of ionization across the lines of force is given by $\mathbf{E}\times\mathbf{B}/B^2$ where $\mathbf{B}$ is the magnetic induction of the earth's field. Motions of ionization along the lines of force are not considered in this paper.

In addition to the tidal wind system there probably is a less obvious wind system caused by the rotation of the neutral atmosphere with the earth at dynamo region heights (about 90–150 km). The geometry of this wind system would be such that the electric field induced in the moving air could be balanced exactly by a polarization field and therefore no ionospheric currents would flow. The polarization field would again cause a drift velocity $\mathbf{E}\times\mathbf{B}/B^2$ which would ensure the rigid rotation of the magnetosphere with the earth.

*It also follows from this condition that the potential distribution over the shell must be symmetrical about the magnetic equator; the symmetry is maintained with the aid of magnetospheric currents along the field lines between the northern and southern hemispheres. In past treatments of the dynamo theory these currents vanished on account of an assumed symmetry about the magnetic equator. Similar assumptions of symmetry are made in the present paper.

It follows from the type of arguments used in the dynamo theory that if the neutral air in the dynamo region were stationary then no motions of the low energy component across the lines of force could occur. These arguments are only approximately correct since they neglect magnetospheric currents perpendicular to the lines of force. However, it is possible to show that for the low pressures of thermal ionization these perpendicular currents are very weak and the motions caused by them are insignificant.

The situation changes completely when energetic particles are present. Even if the density of these energetic particles were considerably lower than the density of the low energy component, currents of the order of a few million amperes could flow perpendicularly to the field at a distance of a few earth radii. The presence of such currents is the basis of Singer's (1957) explanation of the main phase.

A steady ring current caused by energetic trapped particles, whose distribution is symmetrical about the earth's magnetic axis, is free of divergence and therefore does not lead to ionospheric currents. The next section will show how a lack of such symmetry leads to ionospheric currents and magnetospheric motions.

3. THE EFFECTS OF ASYMMETRIC TRAPPING OF PARTICLES

It is convenient to start with the consideration of a relatively simple type of asymmetry in which a new group of trapped particles is introduced predominantly on the day side of the magnetosphere. For simplicity the resulting ring current is assumed to be a line current in the equatorial plane. The value of the current is taken to be greater on the day side than on the night side with abrupt changes of current in the sunrise–sunset plane.

In the subsequent discussions the ring current is taken to be caused entirely by particle drift. There are additional currents due to the gyration of particles but they will be ignored because they are free of divergence.

The use of a line current instead of a current with a finite and large cross section, and the assumption of abrupt rather than gradual changes, are idealizations which simplify the illustration of the mechanism without changing its essential nature. The assumption of initial trapping on the day side is not justified here theoretically; intuitively it appears reasonable. The same assumption has been made by Chamberlain *et al.* (1960), who, however, regard subsequent departures from neutrality as unimportant. Such departures from neutrality are essential to the present mechanism.

Since an axially symmetrical ring current does not give rise to ionospheric currents, it is sufficient to consider only the excess ring current i_R for the calculation of the initial ionospheric current system. This excess current i_R flows along a half circle R in the equatorial plane on the day side, from sunset to sunrise, as illustrated in Fig. 1. In the sunset–sunrise meridian plane the low energy component forms highly conducting paths along the lines of force F, which connect the equatorial current i_R with the polar regions. The circuit is then completed by two ionospheric currents $i_R/2$ in the dynamo region, which flow from A to B in the northern and from C to D in the southern

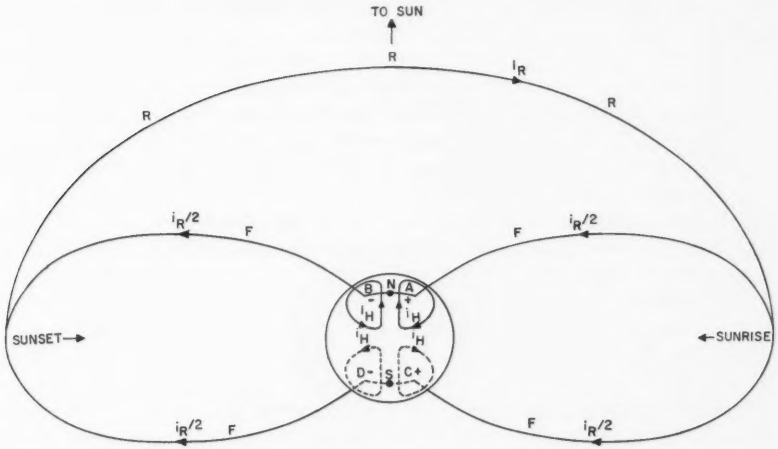


FIG. 1. Perspective view of magnetospheric and ionospheric currents associated with an excess ring current i_R on the day side.

hemisphere. These ionospheric currents are schematically indicated by single flow lines which pass through the north pole N and the south pole S respectively.

The divergence of the current system flowing in the ionospheric shell is thus known. In the idealized current system of Fig. 1 the divergences are at points A and B in the northern and C and D in the southern hemisphere. The ionospheric currents are driven by a quasi-static electric field which has polarization charges as its sources and which is derivable from a potential. The determination of such a current system of known divergence, driven by an irrotational electric field, is a problem well known from the dynamo theory where it has been attacked by numerical methods (for example, Fejer 1953, Maeda 1955).

With the simple geometry of Fig. 1, and if it is assumed that the conductivity of the ionospheric shell does not vary much with geographical position over the polar regions, the general nature of the resulting ionospheric current system may be obtained from relatively simple considerations.

There are two types of current in the ionospheric shell. The first type is *parallel* to the horizontal component of the electric field while the second type, the Hall current, is *perpendicular* to it. If the conductivity of the ionospheric shell is approximately independent of position then the Hall current is almost free of divergence, and the divergence of the parallel current is proportional to the divergence of the horizontal component of the electric field. Thus the horizontal component of the electric field will have a positive divergence, a source at A and a sink at B, in the northern hemisphere; in the southern hemisphere C is the source and D is the sink. The Hall current i_H caused by such an electric field flows in loops around the points A, B, C, and D as

indicated in Fig. 1. On account of the nature of ionospheric conductivity (for example, Baker and Martyn 1953) the Hall current i_H is considerably greater than the current $i_R/2$ which flows parallel to the electric field and completes the circuit of the magnetospheric currents. (The ionospheric current $i_R/2$ flows in directions perpendicular to the Hall current loops. For simplicity only a single flow line of this current is indicated for each hemisphere; likewise only one of the many Hall current loops around each of the points A, B, C, and D is shown.)

The Hall current loops of Fig. 1 show some resemblance to the *Ds* current system of Chapman (1935) and to the current systems derived from magnetic bays by Silsbee and Vestine (1942).

The ionospheric currents described above are driven by a polarization field **E**. The same polarization field also causes motions of the magnetosphere at a velocity $\mathbf{E}\times\mathbf{B}/B^2$. The Hall current results from the motion of the electrons alone at this velocity in the dynamo region, where the motion of ions is stopped by collisions while the electrons are still free to drift. A picture of the motions of ionization above the dynamo region is thus obtained by simply reversing the arrows on the current loops in Fig. 1. The magnetic field lines extend the motions into the magnetosphere in the previously described manner.

The system of currents and motions of Fig. 1 only represents the initial effect of the excess of energetic particles on the day side. The energetic particles, both newly introduced and previously present, are themselves affected by the electric fields which cause the magnetospheric motions. The effect on the previously present belt of energetic trapped particles is particularly important since the ring current carried by it is likely to be very much greater than the excess current i_R of Fig. 1. The approximate effect is taken to be a displacement of the center C of the belt towards the night side of the magnetic north pole N as shown by Fig. 2. Figure 2 is a view of the earth from above the north magnetic pole. The position of the belt of trapped particles is indicated by projection onto the northern hemisphere along the lines of force. The shaded area between the two concentric circles represents the displaced belt. It is assumed for simplicity that such a displacement occurs first and that it is then followed by the usual equatorial drift of the particles of the belt; in reality the two processes occur simultaneously.

The arguments leading to the currents and motions of Fig. 1 can now be used again for the rather different type of asymmetry represented by Fig. 2. In the magnetosphere drifts of the ring current type, indicated by the lines with arrows (which are arcs of circles with their center at N) marked + for protons and − for electrons, are not free of divergence. Currents flow down the lines of force to the dynamo region along the edges of the belt marked with the + signs. The currents then return along other lines of force along the edges marked with the − signs. By previous arguments the same + and − signs represent the divergence of the horizontal component of the electric polarization field on the ionospheric shell. The currents parallel to the electric field flow from the + to the − signs while the Hall currents flow along the lines of equipotential on the shell. Only the Hall currents, which are more intense than the parallel currents, are indicated schematically in Fig. 2. It

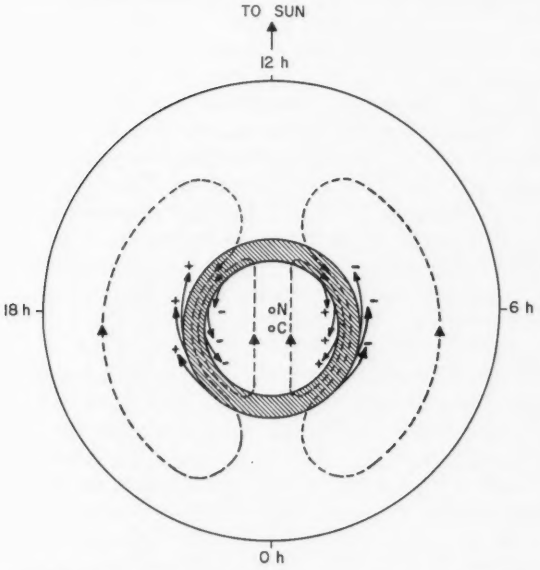


FIG. 2. The ionospheric Hall currents (interrupted lines) caused by a temporary displacement of the trapped particle belt, as seen from above the magnetic north pole N. The projection of the belt along the magnetic lines of force onto the northern hemisphere is indicated by the shaded area between the two concentric circles. The center C of these circles is displaced from N towards the night side.

is to be noted that the electric field, determined by the divergences shown in Fig. 2 by the + and − signs, must be particularly intense in a narrow ionospheric zone which is obtained by projection of the particle belt along the field lines. Intense Hall currents must flow in the same zone.

There will be again motions of ionization in the magnetosphere at the velocity $\mathbf{E}\times\mathbf{B}/B^2$. A picture of the motions immediately above the dynamo region may again be obtained by reversing the arrows on the Hall current loops. The velocity of the motions will be greatest above the same zones where the Hall currents are most intense.

The current system of Fig. 2 resembles the *Ds* current system even more closely than the current system of Fig. 1 because it predicts intense currents in relatively narrow zones (the auroral electrojets) and because it predicts current loops of different sense for high and low latitudes.

A consideration of Fig. 2 should make it clear that the previous presence of several narrow belts of trapped particles would lead to a considerably more complex current system, particularly near the projection of the belts where complete reversals in the direction of the currents could occur within relatively short distances.

The interpretation given in the preceding discussion to the current system of Fig. 2 implies that the resultant current system is obtained by some sort

of superposition of Fig. 1 and Fig. 2. The arguments used in the derivation of these current systems were semiquantitative in nature. A more quantitative theory would have to make more precise assumptions about the energy spectrum and spatial distribution of the previously present and of the newly introduced energetic trapped particles as well as about ionospheric electron densities and collision frequencies. The theory would have to describe the subsequent motion of both the thermal and the energetic components of the magnetosphere until equilibrium is again established largely by the ring current type of drift of energetic particles with a spread of drift velocities appropriate to their energy spectrum.

If the energy of the electrons were predominant then an eastward rotation of the initial current system could be expected with the subsequent weakening of the currents caused by the spread in drift velocities.

If the energy of the positively charged particles were predominant then the rotation would be westward as observed from a co-ordinate system which rotates with the earth. If, however, a co-ordinate system, whose orientation remains fixed with respect to the sun, were used then the rotation of the current system could be either eastward or westward depending on whether the eastward velocity produced by the earth's rotation (by means of the electric polarization field discussed in Section 2) is greater or smaller than the ring current type westward drift velocity. If the ring current type westward drift velocities of the positive energetic particles were about equal in magnitude to their eastward velocity caused by the earth's rotation then the current pattern would hardly rotate at all and would merely decay on account of the spread in drift velocities.

4. THE EFFECTS OF AN ASYMMETRIC CHANGE IN THE EARTH'S MAGNETIC FIELD

An asymmetry similar to that of Fig. 2 could be produced in an entirely different manner through a relatively sudden compression of the earth's magnetic field on the day side by a solar stream as originally suggested by Chapman and Ferraro (1931) in their explanation of the initial phase of a magnetic storm. The compression is assumed to occur in a time much longer than the gyration periods of the particles but much shorter than the time required by the particles to go once around the earth by their ring current type of drift motion.

The result of such a compression of the field has been derived by Parker (1960) whose Fig. 1 resembles the present Fig. 2. If the asymmetry of Fig. 2 were produced in the manner suggested by Parker, then the mechanism of the present paper would not require the introduction of any additional trapped particles. The resulting initial current system would then be that of Fig. 2 alone.

5. GEOPHYSICAL CONSEQUENCES OF MAGNETOSPHERIC MOTIONS

The motions following from the mechanism of the present paper are similar to those of Axford and Hines (1961). Their discussion of some of the consequences of these magnetospheric motions will not be repeated here in detail.

The motions observed in visual and radio aurorae are perhaps the most direct of these consequences.

A more indirect effect of the motions is the acceleration process by 'adiabatic compression', which occurs on the night side and was put forward by Axford and Hines as a possible basis for an auroral theory. An equivalent explanation of this acceleration process has been given in their paper in terms of the motion of individual energetic charged particles.

This acceleration process on the night side with the resulting lowering of mirror points and eventual loss of trapped energetic protons and electrons (at different latitudes) could lead to an additional current system. The formation of a current system in a similar manner, but by a different acceleration process, has been suggested by Chamberlain *et al.* (1960) and Kern (1961).

6. CONCLUSIONS

It was attempted to show in the present paper that the assumption of certain types of temporary asymmetry in the distribution of trapped particles relative to the magnetic field leads to a system of currents, motions, and acceleration processes which appear to explain, at least qualitatively, many of the observed phenomena associated with magnetic disturbances and aurorae.

ACKNOWLEDGMENTS

The author is indebted to Prof. T. Gold and Drs. C. O. Hines and W. I. Axford for drawing the author's attention to the problem of magnetospheric motions and for interesting discussions. The work was performed under Project No. D48-95-10-27.

REFERENCES

AXFORD, W. I. and HINES, C. O. 1961. Can. J. Phys. **39**. This issue.
BAKER, W. G. and MARTYN, D. F. 1953. Phil. Trans. Roy. Soc. A, **246**, 283.
CHAPMAN, S. 1935. Terrestrial Magnetism and Atmospheric Elec. **40**, 349.
CHAPMAN, S. and FERRARO, V. C. A. 1931. Terrestrial Magnetism and Atmospheric Elec. **36**, 77, 171.
CHAMBERLAIN, J. W., KERN, J., and VESTINE, E. H. 1960. J. Geophys. Research, **65**, 2535.
DESSLER, A. J. and PARKER, E. N. 1959. J. Geophys. Research, **64**, 2239.
FEJER, J. A. 1953. J. Atmospheric and Terrest. Phys. **4**, 184.
GOLD, T. 1959. J. Geophys. Research, **64**, 1219.
KERN, J. W. 1961. J. Geophys. Research, **66**, 1290.
MAEDA, H. 1955. J. Geomagnetism and Geoelec. **7**, 121.
PARKER, E. N. 1960. J. Geophys. Research, **65**, 3117.
POPE, J. H. 1961. J. Geophys. Research, **66**, 67.
SCHMIDT, A. 1924. Z. Geophys. **1**, 1.
SILSBEE, H. C. and VESTINE, E. H. 1942. Terrestrial Magnetism and Atmospheric Elec. **47**, 195.
SINGER, S. F. 1957. Trans. Am. Geophys. Union, **38**, 175.
SMITH, E. J., COLEMAN, P. J., JUDGE, D. L., and SONETT, C. P. 1960. J. Geophys. Research, **65**, 1858.
STORMER, C. 1911. Arch. sci. phys. et nat. **32**, 117, 190, 277, 415, 501.
VAN ALLEN, J. A. and FRANK, L. A. 1959. Nature, **183**, 430.

TRANSITION RANGES FOR THREE-DIMENSIONAL WAKES[1]

R. H. MAGARVEY AND ROY L. BISHOP

ABSTRACT

The motions set up in a liquid due to the passage of a drop of an immiscible liquid have been examined for a number of liquid–liquid systems. Wake configurations corresponding to Reynolds numbers between 0 and 2500 have been photographed, and the transition ranges between the various wake configurations have been determined. Several liquid–liquid systems were used in obtaining data on wake patterns, and it was found that identical patterns correspond to approximately the same Reynolds numbers regardless of the system. However, the exact transition number cannot be determined for the general case as drop deformations depend on the physical characteristics of a particular system. The range of Reynolds numbers corresponding to a transition from one wake pattern to another is somewhat less than 20.

INTRODUCTION

Apart from practical considerations, wakes in liquid–liquid systems are interesting physical phenomena in their own right. The periodicity and symmetry evident in some of the three-dimensional wake forms are comparable with any associated with two-dimensional patterns. Three-dimensional wakes corresponding to drop motions characterized by Reynolds numbers less than about 210 are single threads, and those corresponding to numbers between 290 and 700 consist of a double row of vortex rings. This double row configuration has been photographed by Magarvey and Bishop (1960, 1961).

The Reynolds number is the most convenient flow index by which wake configurations associated with liquid–liquid systems can be compared, but due to the deformable nature of the drop the precision associated with solid–liquid systems is lacking. The Reynolds number is given by

$$Re = LV/\nu \tag{1}$$

where L is some characteristic length, V is the relative velocity between the drop and the quiescent phase, and ν is the kinematic viscosity of the disperse liquid. In these experiments, the major axis of the deformed drop was found to be a more pertinent dimension than the equivalent diameter.

As two-dimensional wakes have been investigated for a wide range of Reynolds numbers, a brief examination of the phenomena is indigenous to a consideration of the three-dimensional counterparts. Essentially, with increasing Reynolds numbers, an increasing amount of vorticity originating in the boundary layer is concentrated into two eddies on the immediate downstream side of the obstacle. The vortices thus formed gain strength, increase in size, and change shape as the Reynolds number is increased. Eventually the standing eddies become so elongated that they break down, and, in the case of a cylinder, a characteristic state of flow is developed in which vortices are

[1]Manuscript received June 7, 1961.
Contribution from the Department of Physics, Acadia University, Nova Scotia. The research for this paper was supported in part by the Defence Research Board of Canada, under grant number D.R.B. 9550-04 and in part by the Nova Scotia Research Foundation.

shed alternately from the two sides of the obstacle. The eddies are shed at regular intervals, forming a wake configuration called a Kármán vortex street. In general, the Reynolds number at which the shedding is initiated depends on the shape of the obstacle. The frequency of shedding is given in terms of a dimensionless parameter called the Strouhal number. This number is given by

$$S = fL/V \tag{2}$$

where f is the frequency, V the velocity, and L the same dimension as in equation (1). Rayleigh (1915) obtained an empirical relationship between the Strouhal and Reynolds numbers for flow past circular cylinders. The equation relating these two dimensionless groups and the range for which they are valid are indicated by

$$S = 0.195\left(1-\frac{20.1}{Re}\right) \qquad 40 < Re < 1000. \tag{3}$$

The theoretical three-dimensional generalization of the standing eddy phenomenon is a ring with the axis coinciding with the direction of flow. This ring has been observed on the downstream side of flat plates and bodies of revolution for a limited range of Reynolds numbers. Stanton and Marshall (1930) observed vortex rings for very small Reynolds numbers behind circular plates. The nature of the wake outside the vortex sheet enclosing the two-dimensional eddy pair or the three-dimensional ring has received little attention other than for cases of periodic shedding. This report concerns three-dimensional wakes corresponding to a range of Reynolds numbers extending from 0 to about 2500. The range of numbers over which the transition from one wake pattern to another occurs has been found experimentally to be less than 20.

WAKE CLASSIFICATION

It is convenient to classify wakes according to the nature of the motions set up in the continuous phase far behind the moving body. Although the motions within the vortex sheet on the downstream side of the body determine the nature of the trail far behind, these motions are much more difficult to observe. In liquid–liquid systems, six distinct and reproducible trail configurations are recognizable for the entire range of stable drops. The range of numbers corresponding to each classification is indicated in Table I. The line

TABLE I

Class	Range of Reynolds numbers	Nature of trail
I	0–210	Single thread
II	210–270	Double thread
III	270–290	Double thread with waves
IV	290–410	Procession of vortex loops
V	290–700	Double row of vortex rings
VI	700–2500	Asymmetrical wake

of demarcation is not sharp between any two classes as the values are a result of observations in different systems. The identical lower limit for classes IV and V cannot be explained from an examination of the photographic evidence.

STEADY WAKES

Two flow types are associated with the single thread wake. For very small Reynolds numbers the streamlines follow the interface and exhibit a symmetry for spheres and other symmetrical forms. For Reynolds numbers above a critical value, the streamlines separate from the surface and form a closed region immediately behind the moving body. The exact transition Reynolds number has not been determined, but experimental evidence indicates a value less than 10. The trail left by the passing drop is a single thread regardless of separation, provided the Reynolds number is less than 210. The vortex layer which leaves the interface forms a vortex sheet separating the wake liquid from the main flow. Inside the sheet an inflow along the axis of the wake gives rise to a vortex ring immediately behind the drop. The single thread is sufficient to maintain an equilibrium between the rate at which vorticity is generated and the rate at which it is dispensed into the main stream. Figures 1(*a*) and (*b*) show the single thread some distance behind the advancing drop and a section of the ring inside the sheet.

One of the most interesting wake configurations is the double thread observed behind drops characterized by Reynolds numbers between 210 and 270. An asymmetry in the circulatory motion within the vortex sheet produces a corresponding asymmetry in the sheet itself and a resultant deviation from a vertical line of fall. As no rotation is observed, it is assumed that the lack of symmetry is responsible for the sidewise force component. The exact nature and cause of this lift cannot be inferred from the photographic data. The double thread maintains an equilibrium between rates of generation and diffusion of vorticity.

Gunn (1949) observed deviations from the vertical direction of fall for water drops of the same Reynolds numbers as characterize these class II drops in liquid–liquid systems. The water drops were falling through air, and Gunn attributed the behavior to a mechanical resonance phenomenon associated with eddy detachment and a natural frequency of drop oscillation. Since a comparable motion can be observed directly in liquid–liquid systems, it is entirely possible that the two phenomena are identical. However, there are definitely no aeolian sheddings and no oscillations associated with liquid–liquid systems in this range of Reynolds numbers. Figures 2(*a*) and (*b*) indicate the wakes far behind and close to the moving drop. The circulations within the vortex sheet are in evidence in Fig. 2(*b*). The transition Reynolds number from a single to a double thread configuration is fairly sharp for a specific system, but for the general case, 210 is the approximate value.

PERIODIC WAKES

The state of equilibrium set up under conditions corresponding to the first two classes is disturbed by increasing further the Reynolds number. The rate

FIG. 1. (*a*) and (*b*) The wake behind a carbon tetrachloride drop moving through water at terminal velocity. Reynolds number 170.

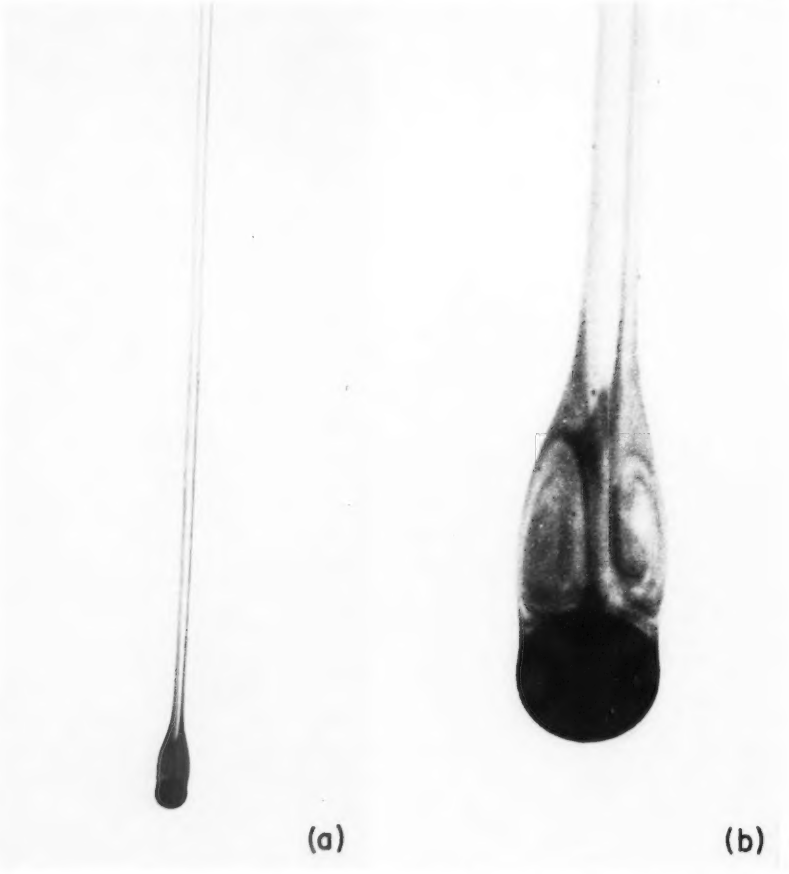


FIG. 2. *(a)* and *(b)* Double thread wakes for two different drops. The dye concentration indicates the pattern of circulation within the vortex sheet. Reynolds numbers 225 and 220 respectively.

FIG. 3. (*a*) and (*b*) Wavy double thread wake behind drop of monochlorobenzene moving through water.

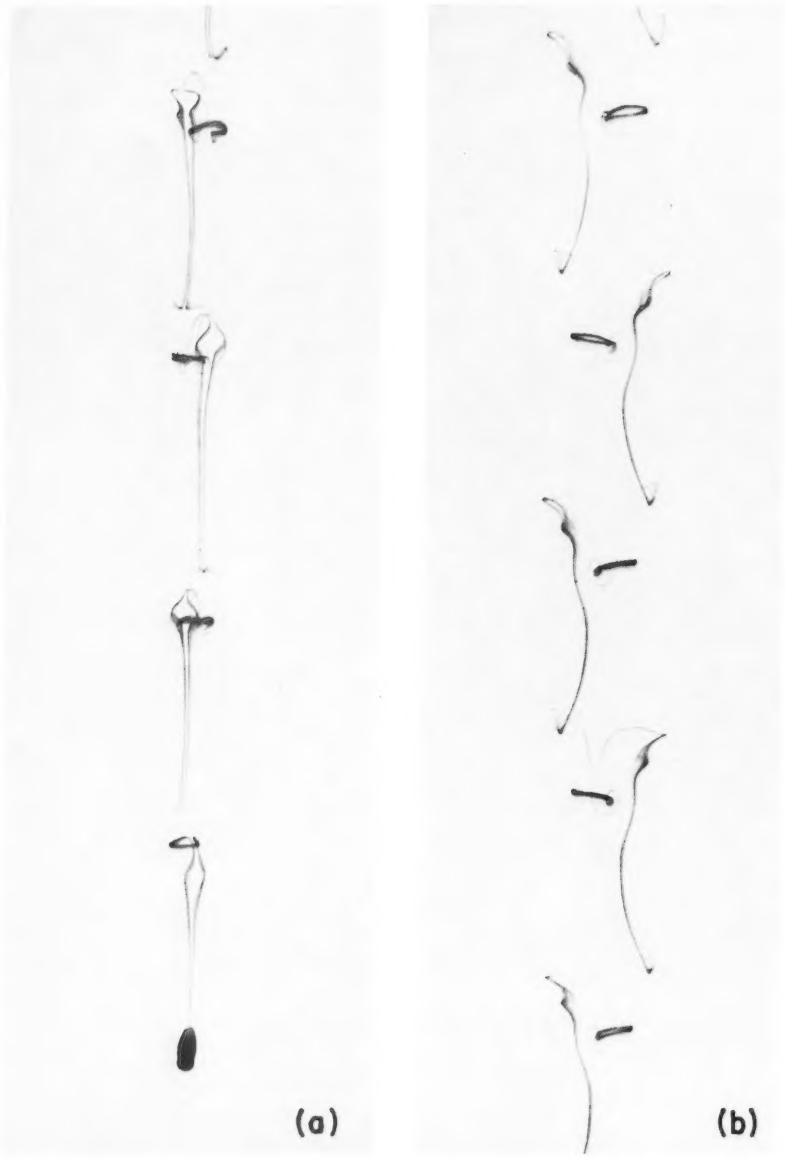


FIG. 4. (*a*) and (*b*) Two views of class IV loop wake. Interval of 5 seconds between photographs. Reynolds number 370.

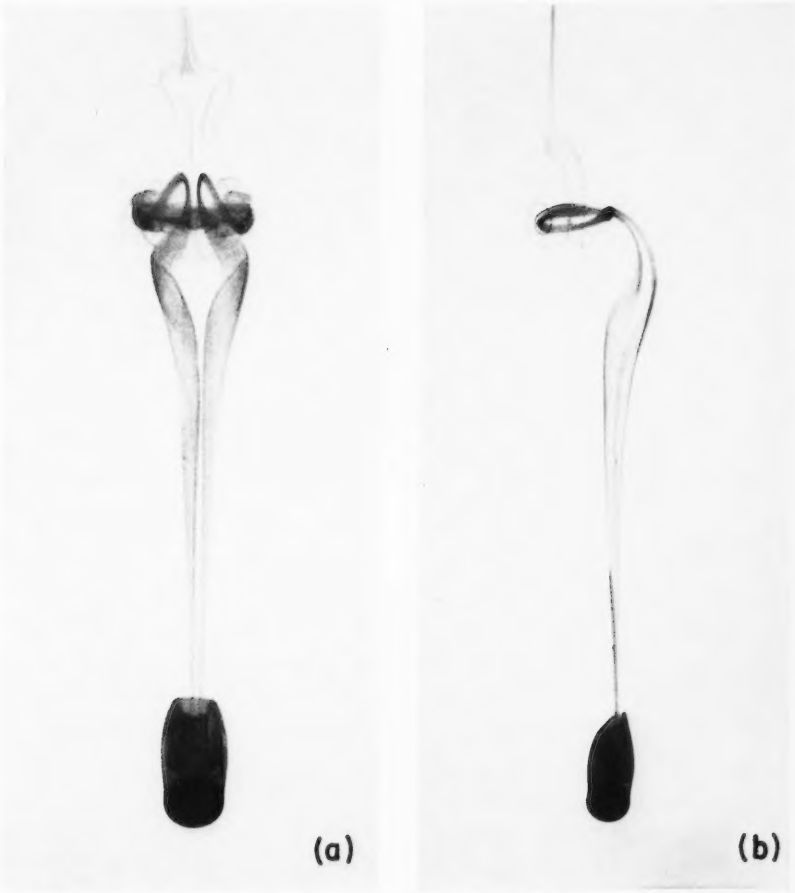


FIG. 5. (*a*) and (*b*) Two views of single cycle of a loop wake. Circulation inside vortex sheet shown from two angles.

FIG. 6. Class V wake consisting of double row of vortex rings. Reynolds number 360.
FIG. 7. Rings associated with class VI wake. Reynolds number 1040.

at which vorticity is diffused from the sheet into the main body of the fluid remains practically constant, but the increased rate at which it is transferred to the vortex ring creates unstable conditions within the vortex sheet. The release of vorticity at regular intervals is the first suggestion of periodic wakes. Basically, the process is one of build-up and release, but no sizeable portion of the ring escapes through an opening in the end of the vortex sheet during the cycle. When the vortex strength of the ring reaches a critical value, a sudden motion of the ring disturbs the sheet, which, in turn, is responsible for a release of vorticity and a consequent return of the ring to its original position and shape. The class III wake corresponds to a narrower range of Reynolds numbers than the first two, and could be considered as a transition range. However, according to the system of classification based on the nature of the trail far behind the drop, this type of wake must be considered as a separate class. The wake appears as a double thread characterized by equally spaced, wavy disturbances. Figures 3(*a*) and (*b*) indicate the wavy trail and the circulation pattern immediately behind the drop.

For Reynolds numbers greater than 290, the trail assumes a repetitive pattern and the circulation within the sheet ceases to resemble a symmetrical vortex ring. Two distinct wake configurations correspond to the range of Reynolds numbers extending from 290 to 410: the first is characterized by a series of vortex loops connected by an intricate system of vortex filaments, and the second is the double row of vortex rings described by Magarvey and Bishop. The two wake types have a common lower limit, but the symmetrical vortex ring pattern extends to an upper limit of 700. The occurrence of either the loop or the ring wake in the overlapping region of Reynolds numbers appears to be a matter of chance. Initial conditions of drop formation probably are responsible for the flow which determines the nature of the wake.

In the cycle of build-up and release, the vorticity generated in the boundary layer is concentrated on diametrically opposite sides of the flow axis within the vortex sheet. The sections in which the vortex strength is greatest are alternately discharged into the main body of the fluid. With each ejection a portion of the sheet is carried away. The vortex element discharged into the stream interacts with the disperse liquid to form a regular wake pattern. The vortex loop configuration is termed a class IV wake and the vortex ring street a class V wake. Figures 4(*a*) and (*b*) show two separate loop wakes photographed at different angles. Figures 5(*a*) and (*b*) show the circulation patterns within the vortex sheet for this type of wake.

An examination of many loop wakes indicates that the frequency of discharge is a function of velocity and the physical characteristics of the disperse phase. The physical properties of the drop phase are of no less importance, but they appear as factors in determining the velocity of the drop. Figure 8 shows the Strouhal number plotted against the Reynolds number for wakes over the entire range of Reynolds numbers for which class IV wakes are observed. As the best fitting curve is a straight line, the relationship between the frequency and velocity is given in terms of the slope.

$$f = (V^2/\nu) \text{ slope} \tag{4}$$

where ν is the kinematic viscosity of the field phase. Magarvey and Bishop (1961) found a linear relationship between the Strouhal and Reynolds numbers for the range corresponding to class V wakes. Möller (1938) made a comparable plot for rigid spheres covering a range of Reynolds numbers for which class VI wakes are observed in liquid–liquid systems. However, the wakes behind the spheres investigated by Möller do not exhibit the periodicity of the class IV wakes behind drops.

Figure 6 shows a class V wake photographed in a carbon tetrachloride–water system. The frequency of discharge is greater than those of class IV.

The symmetry and periodicity characterizing the wakes of classes IV and V are not in evidence for Reynolds numbers greater than 700. Vortex rings are still the main features of these wakes, but there appears to be no preferred plane at which the individual elements are discharged into the quiescent phase. The vorticity generated in the boundary layer is concentrated into the liquid within the sheet at such a rate that no definite pattern of circulation can be established before an element is discharged. The discharge of one element does not initiate conditions which lead to a discharge from the diametrically opposite side of the wake axis. Figure 7 shows a typical class VI wake.

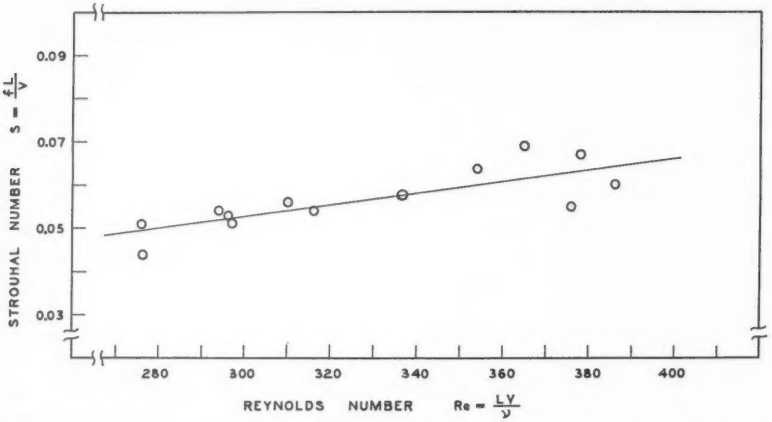


FIG. 8. Graph showing relationship of Strouhal and Reynolds numbers for loop wakes

REFERENCES

GUNN, R. 1949. J. Geophys. Research, **54**, 383.
MAGARVEY, R. H. and BISHOP, R. L. 1960. Nature, **188**, 735.
——— 1961. Phys. Fluids, **4**, 800.
MÖLLER, VON W. 1938. Physik. Z. **39**, 37.
STRUTT, J. W. (LORD RAYLEIGH). 1915. Phil. Mag. (6), **29**, 453; Sci. Papers, **6**, 315.
STANTON, SIR T. E. and MARSHALL, D. 1930. Proc. Roy. Soc. A, **130**, 295.

ON THE STATISTICAL PROPERTIES OF OPTICAL AND RADIO AURORA[1]

F. Hector and P. A. Forsyth

ABSTRACT

The latitudinal variation of certain auroral forms has been investigated by means of all-sky camera records made in central Canada during the I.G.Y. The forms chosen were those that may have significance in theories of auroral morphology. The occurrence of "curl forms" seems to be concentrated near the auroral zone. One type of radio reflection (A_2) appears to be associated with the curl forms and another type (A_3) with patchy, early morning aurora, which occurs most frequently well south of the auroral zone.

INTRODUCTION

The remarkable similarity between certain auroral configurations and the instability patterns that develop in a sheet beam of charged particles moving along a magnetic field, seems first to have been brought out in discussions between Webster and Kavadas (see Webster 1957). The auroral configuration of interest is the so-called "curl form" in which a section of an arc or band develops a kink which rapidly progresses into an "*S*" curve (as seen in plan view) and thence into a nearly isolated spiral configuration. The curl form appears to represent the transition between the quiet arc form and the broken ray forms. If the obvious similarity between this auroral sequence and the break-up of sheet beams is more than superficial, it implies that the auroral configurations are actually determined by beams of incoming particles that travel considerable distances. The alternative view is that the shapes of auroral forms are determined by local conditions in the ionosphere.

Recently, Kern and Vestine (1961) have examined several features of auroral morphology, including motions in auroral forms, on the assumption that sheet beams are involved and that the sheet beams are made up of particles trapped in the geomagnetic field. Earlier, Gartlein (1959) pointed out that if the auroral configurations are governed by the precipitation of geomagnetically trapped particles, the occurrence of curl forms should vary with geomagnetic latitude, because of the varying length of the geomagnetic field lines. It might be expected that the relative occurrence of curl forms should increase with increasing geomagnetic latitude, at least up to some latitude where the auroral arcs cease to occur.

In the course of a radar investigation of aurora, Lyon (1960) found a notable similarity between the diurnal variation of curl forms and the occurrence of radar echoes for a particular location (near Flin Flon). In discussing this result, Forsyth (1960) drew attention to the particular geometry involved which made it uncertain whether the association of curls with radar echoes

[1]Manuscript received June 26, 1961.
Contribution from the Institute of Upper Atmospheric Physics, University of Saskatchewan, Saskatoon, Sask.

was due to an increase of ionization when the curls are formed or to a kind of observational selection that was peculiar to the radar situation. It was suggested that this point could be resolved by means of continuous-wave, bistatic, radio observations in the same general region.

This paper attempts to shed some light on these two characteristics of curl forms: the variation of occurrence with latitude and the association with radio reflections. In the course of the study interesting characteristics were found also for the patchy, early-morning aurora. The results were obtained by analysis of all-sky camera films and bistatic radio records obtained during the International Geophysical Year.

RESULTS AND ANALYSIS

During the I.G.Y. all-sky auroral cameras were operated at Uranium City (59.5° N., 108.6° W.), Flin Flon (54.8° N., 101.8° W.), Saskatoon (52.1° N., 106.6° W.), and Regina (50.4° N., 104.7° W.). The choice of a suitable period for analysis required a compromise between the need for a statistically acceptable body of data and the requirement that it should be as homogeneous as possible. The vagaries of weather and equipment performance precluded any completely satisfactory solution. The period chosen was 27 November, 1957, to 28 February, 1958. The number of useful records obtained in this period for each station is shown in Table I. The number of hours of clear-

TABLE I

All-sky film records

	No. of hours of useful film	Percentage of dark hours
Uranium City	183	14
Flin Flon	437	34
Saskatoon	286	22
Regina	368	28

weather film obtained, particularly at Uranium City, was a small fraction of the total number of dark hours. However, this situation prevailed throughout the I.G.Y. and there was, in fact, a strong concentration of useful records in the period chosen. It is felt that, while the number of data is less than is desirable, the validity of the analysis is not likely to be improved by the inclusion of additional, much less homogeneous, data from the remainder of the I.G.Y.

The films were analyzed by assigning to each 10-minute period a classification representing the most prominent type of aurora present. The classifications used were:

(i) arcs, including only well formed arcs, each with a clearly defined lower edge without scallops or folds;

(ii) curls, the curl forms described earlier and including arcs with wavy structure along the lower edges as seen in the photographs;

(iii) patchy, including all the forms that appear as luminous patches on the photographs but probably referring mostly to pulsating patches;

(iv) other, all other forms including diffuse surfaces, active rayed bands, and rays.

The results of the analysis are presented in Fig. 1 where the percentage occurrence (percentage of the total number of clear-sky, 10-minute periods)

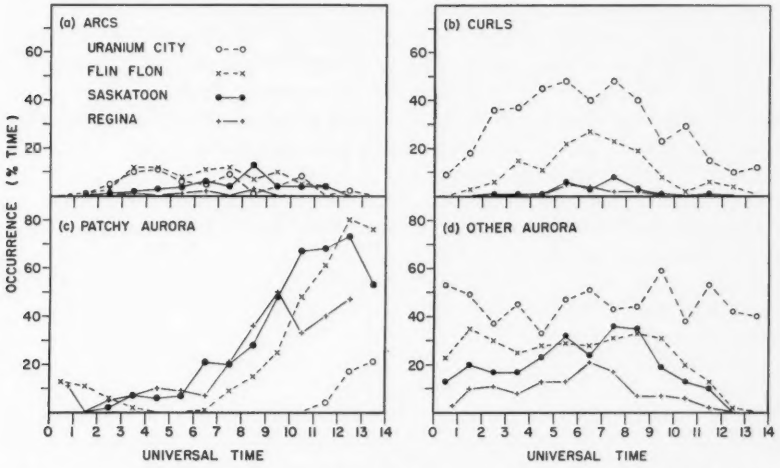


FIG. 1. The variation of occurrence with time of various auroral forms, for the dark hours. For each station, local midnight occurs near 7 hours U.T. (see Table II).

of each classification in each hour of the night is plotted. Except for the first and last hours (0–1 U.T. and 13–14 U.T.) there was a fairly uniform distribution of useful film records over the night hours for each station.

Bistatic Radio Observations

The radio measurements made use of two of the bistatic paths, Yellowknife–Saskatoon and Yellowknife–The Pas, that are described elsewhere (Forsyth, Green, and Mah 1960). As may be seen from Table II, the geomagnetic latitude

TABLE II

Geomagnetic latitude and local time

	Geomagnetic latitude	Time of local midnight (U.T.)
Uranium City	67.4°	7 hr 14 min
Flin Flon	63.8°	6 hr 48 min
Saskatoon	60.5°	7 hr 6 min
Regina	59.1°	6 hr 59 min
Yellowknife–The Pas (mid-point)	66.6°	7 hr 8 min
Yellowknife–Saskatoon (mid-point)	64.9°	7 hr 20 min

of the mid-point of the Yellowknife–Saskatoon path is slightly greater than that of Flin Flon. The mid-point of the Yellowknife–The Pas path is near, but south of, Uranium City. The records were the same as those used earlier,

but this time they were subjected to a more detailed analysis. As before, the record interval used for analysis corresponded to 15 minutes in time. For each interval the presence of an enhanced signal, due to scattering of the very high frequency radio waves by auroral ionization was listed as an event. The earlier study made use only of A_2 and S events. The present study included also the events classified as A_1 and A_3. The classifications used here (A_1, A_2, and A_3) were described and illustrated in an earlier paper (Collins and Forsyth 1959). Although the A_3 events were of low amplitude on the records it was found that with care they could be distinguished with reasonable certainty. Since the analysis was confined to the night hours, S events were not encountered. As indicated in both of the earlier studies, the occurrence of A_1 events at these latitudes is negligibly small and will not be considered further.

The radio records were independent of weather conditions and, except for occasional periods of equipment failure, were continuous throughout the 3-month period. In all, the number of hours for which useful records were obtained amounted to 1200 for the Yellowknife–Saskatoon path and 1233 for the Yellowknife–The Pas path. The results of the analysis are presented in the next section.

DISCUSSION

It is unfortunate that the all-sky camera results from Uranium City are based on so few data. The very high percentage occurrence of curl forms would seem to offer strong support to Gartlein's contention that the occurrence of these forms should increase with increasing geomagnetic latitude. While the details of the hour-to-hour variations for Uranium City, shown in Fig. 1, should be treated with some caution it seems clear that curl forms did occur at Uranium City in substantially greater numbers than at Flin Flon. There was also a tendency for these curl forms to be concentrated near the middle of the night.

In contrast to those from Uranium City, the films from Flin Flon and Regina were excellent and those from Saskatoon were acceptable. The graphs of Fig. 1(*a*), (*b*), and (*d*) show, quite clearly, a marked decrease of occurrence of aurora with decreasing latitude. The percentage occurrence of all aurora at each station was: Uranium City, 83%; Flin Flon, 65%; Saskatoon, 52%; and Regina, 30%. Even more interesting is the sharp *relative* decrease of occurrence of the curl forms. It would appear that the curl forms only occur in significant numbers in or near the auroral zone, a conclusion that is in accord with the implications of the sheet-beam mechanism of curl formation discussed by Gartlein.

Another interesting feature of Fig. 1 is the extremely frequent occurrence of the "patchy" aurora in the morning hours. While the occurrence of this type of aurora has often been noticed by visual observers, it seems to have received little attention. Heppner (1954) showed that it was primarily an early-morning phenomenon and that it seemed to occur most frequently near geomagnetic latitude 64° (well south of the maximum of the auroral zone). Both these conclusions are supported by the present results.

An unexpected feature of Fig. 1 is the relative unimportance of the arcs (Fig. 1(*a*)). This result is not in agreement with that of Lyon, who found a high peak in the occurrence of quiet arcs during the early part of the night. The difference between the two results probably is due to the more stringent criteria used in the present study. An examination of Fig. 1(*a*) does show a tendency for the arcs to occur at later times with decreasing latitude. This characteristic also supports the sheet-beam hypothesis.

The results of the analysis of radio records are shown in Fig. 2. Also shown in the same figure are the diurnal variations of curls and patchy aurora at Flin Flon and Uranium City (from Fig. 1(*b*) and (*c*)). In spite of the separations

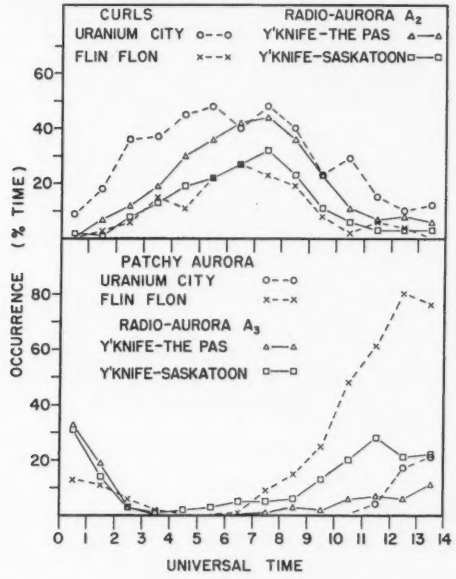


FIG. 2. A comparison of the variation of occurrence with time of curl forms and of one type of radio scattering (A_2), and a similar comparison for patchy aurora and another type of radio scattering (A_3).

between the all-sky cameras and the path mid-points and of the uncertainties in the results from Uranium City, it seems clear that there is a close similarity between the diurnal variation of the curl forms and that of the A_2 events. This confirms the finding of Lyon and indicates that his result was not due to the particular geometry of his experiment. It must be concluded that the observed association between the radio reflections and the curl forms is due to the presence of intense ionization in these structures. In Fig. 2, the radio A_2 events observed on the Yellowknife–Saskatoon path appear to lag slightly in time behind the curl forms at Flin Flon. This lag is probably due to the fact that local time at the radio path mid-point is about half an hour behind local time at Flin Flon (see Table II).

In Fig. 2, there is also some similarity between the variations of the A_3 events and the patchy aurora. The association has often been noticed but this seems to be the first time that statistically significant evidence has been presented. The radio equipment used during the I.G.Y. was considerably less sensitive than that used for other investigations. When more sensitive equipment is used, the A_3 events usually dominate the results. If this had been the case for the I.G.Y. experiment the similarity between the diurnal variation of the patchy aurora and that of the A_3 events would have been even closer. Since the faint, patchy aurora that occurs at the end of an optical display is often disregarded in visual studies, it is not surprising that many observers have concluded that there is little relation between optical and radio aurora.

If the association between the curl forms and the A_2 radio reflections can be accepted, then the radio evidence can be used to expand our knowledge of auroral morphology. There is ample evidence (Forsyth, Green, and Mah 1960; Green 1961) to show that the peak occurrence of A_2 events occurs near midnight at subauroral latitudes and progressively earlier with increasing latitude. From the discussion of Kern and Vestine it would appear that this behavior might be expected of the curl forms, if they do arise as a result of instabilities set up in sheet-like configurations of particles trapped in the geomagnetic field.

ACKNOWLEDGMENTS

The all-sky film records are the property of the Institute of Upper Atmospheric Physics, University of Saskatchewan. The radio records are the property of the Radio Physics Laboratory of the Defence Research Board of Canada. The authors are grateful for the opportunity to use and to intercompare these two sets of records.

REFERENCES

Collins, C. and Forsyth, P. A. 1959. J. Atmospheric and Terrest. Phys. **13**, 315.
Forsyth, P. A. 1960. Can. J. Phys. **38**, 593.
Forsyth, P. A., Green, F. D., and Mah, W. 1960. Can. J. Phys. **38**, 770.
Gartlein, C. W. 1959. Ann. géophys. **15**, 31.
Green, F. D. 1961. Ph.D. Thesis, University of Saskatchewan.
Heppner, J. P. 1954. Ph.D. Thesis, California Institute of Technology, Pasadena, California.
Kern, J. W. and Vestine, E. H. 1961. J. Geophys. Research, **66**, 713.
Lyon, G. F. 1960. Can. J. Phys. **38**, 385.
Webster, H. F. 1957. J. Appl. Phys. **28**, 1388.

THE ELECTRICAL RESISTIVITY OF Mn_3ZnC BETWEEN 4.2° AND 630° K[1]

M. L. SWANSON[2] AND S. A. FRIEDBERG[3]

ABSTRACT

The electrical resistivity of an alloy having approximately the composition Mn_3ZnC has been measured between 4.2° and 630° K. From a high residual value the resistivity increases with temperature in a complicated way up to 550° K, becoming temperature independent above that point. Two anomalous regions are evident corresponding to the magnetic transitions occurring in this substance at ~240° K and ~450° K. The general features of the results are interpreted on the basis of a spin-disorder scattering picture.

INTRODUCTION

Alloys with compositions near Mn_3ZnC exhibit unusual magnetic properties (Butters and Myers 1955). Below a Curie temperature θ_C (~450° K), which varies with exact composition, these substances become ferromagnetic. At a lower temperature, θ_N ~240° K, neutron diffraction (Brockhouse and Myers 1957) and heat capacity (Swanson) data indicate another co-operative transition, the saturation moment at this point passing continuously through a maximum value. On the basis of their neutron diffraction studies, Brockhouse and Myers have concluded that between θ_C and θ_N all Mn atoms in these alloys are characterized by the same moment and that these are ferromagnetically aligned. Below θ_N the observations are described by assuming that the Mn atoms have different moments, 0, 2, or 3 Bohr magnetons, and that a ferrimagnetic arrangement of spins exists.

In view of the redistribution of spins among Mn atoms presumed to accompany the ordering transition at θ_N, we have felt it to be of some interest to examine the temperature variation of electrical resistance in Mn_3ZnC over an extended range including both θ_N and θ_C. Such observations might be expected to suggest what role, if any, the electrons responsible for the magnetization play in the conduction process.

SPECIMEN MATERIAL

An alloy having a composition close to Mn_3ZnC was prepared by sintering Mn_3C and Zn for 8 days at 600° C and then for 2 days at 900° C in an evacuated quartz capsule. The high temperature sinter yielded a brittle alloy of metallic appearance. Debye–Scherrer diffraction photographs showed the alloy to contain approximately 95% of the desired perovskite structure with lattice parameter 3.921_5 Å. The density of the alloy was 6.13 ± 0.06 g/cm^3, as compared

[1]Manuscript received July 3, 1961.
Contribution from the Carnegie Institute of Technology, Pittsburgh, Pa., U.S.A. This research was supported in part by the Office of Naval Research, the National Science Foundation, and the Alfred P. Sloan Foundation.

[2]Metals Research Laboratory. Now at Atomic Energy of Canada Ltd., Chalk River, Ontario, Canada.
[3]Department of Physics.

with the density 6.63 g/cm³ calculated from the lattice parameter. According to Butters and Myers (1955) such a lattice parameter should characterize a material of the composition $Mn_{62.1}Zn_{17.9}C_{20}$ with a Curie temperature of 416° K. Chemical analysis indicated, however, more nearly 62.8 at. % manganese, suggesting that the average lattice parameter of the specimen was lower than that quoted above and that the observed Curie temperature should be higher than 416° K. A crude determination of the Curie temperature indicated, in fact, the value $\theta_C \sim 453°$ K.

EXPERIMENTAL METHOD

The electrical resistivity of Mn_3ZnC was measured with an accuracy of 0.5% by the potentiometric method from 4° to 630° K. Mechanical and soldered contacts gave consistent results. Above the liquid nitrogen range, temperatures were measured to ±1° K with a calibrated copper–constantan thermocouple. Lower temperatures were determined by means of the vapor pressure – temperature relation for the particular refrigerant employed, for example, liquid H_2 or He.

RESULTS AND DISCUSSION

The electrical resistivity of a specimen of Mn_3ZnC is shown as a function of temperature in Fig. 1. The magnitude of the resistivity at 298° K is 370 micro-ohm cm, roughly one-half that reported by Butters and Myers for a loosely sintered alloy of similar composition. When allowance is made for the fact that the present material, while closer in density to the crystal value, is

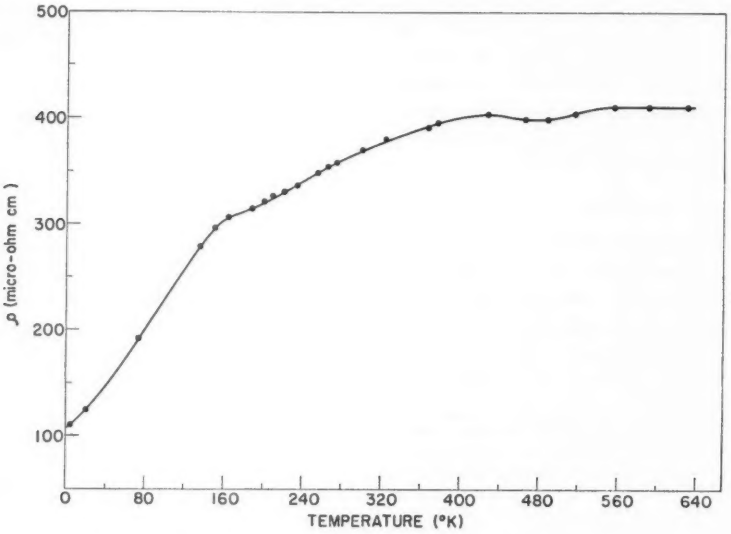


FIG. 1. Resistivity of Mn_3ZnC as a function of temperature.

still not fully compact, its room temperature resistivity is comparable with those of other ferromagnetic metals and alloys. The general downward trend of ρ with temperature toward a residual value of 107 micro-ohm cm at 0° K is further indication of the metallic character of Mn_3ZnC.

Several features of the ρ vs. T curve are noteworthy. First of all, it becomes essentially temperature independent above about 550° K. Secondly, distinct anomalies appear in the curve corresponding to the two magnetic transitions in Mn_3ZnC. The higher of these anomalies in temperature includes a small but reproducible region of negative slope. A plot of the temperature coefficient $(1/\rho)(d\rho/dT)$ vs. T shown in Fig. 2 brings out more clearly the features of the anomalies. The maximum at ~241° K coincides quite closely with the transition temperature θ_N determined calorimetrically by Swanson on a similar

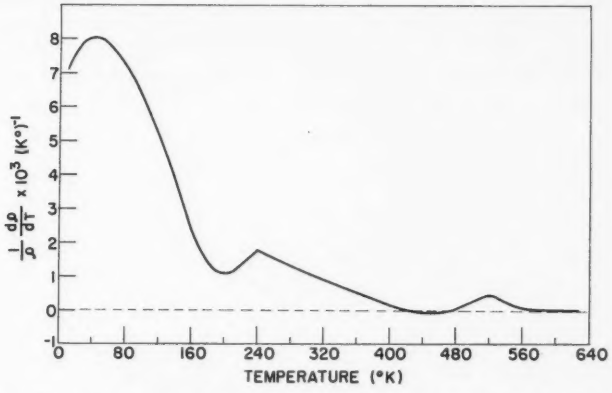


FIG. 2. The temperature coefficient of resistivity $(1/\rho)(d\rho/dT)$ as a function of temperature for Mn_3ZnC.

alloy, $\theta_N = 238°$ K. The negative minimum near 450° K occurs approximately at the Curie point, θ_C, estimated for this alloy from rough magnetic observations. It may be remarked that the range in which this material exhibits a negative temperature coefficient of resistivity is much more restricted than that found by Butters and Myers for their loosely sintered alloy.

The general character of the ρ vs. T curve for Mn_3ZnC just described suggests that an important resistive mechanism in this substance is probably spin-disorder scattering.* This scattering arises from the interaction of the conduction electrons with unpaired electrons assumed to be localized on the magnetic ions and to be largely responsible for the magnetization of the solid. At 0° K the ionic moments are presumed to be perfectly ordered so that no scattering occurs. Disorder among the ionic moments increases with rising temperature, producing a corresponding increase in the spin-disorder contribution to the resistance until the transition temperature is reached. Above the transition temperature long-range order is completely destroyed

*A review of this subject with full references has been given by Coles (1958).

and in the absence of short-range order the spin-disorder resistance would be expected to be temperature independent. DeGennes and Friedel (1958) have shown that short-range effects can cause the resistivity not only to be temperature dependent immediately above the transition point but actually to exhibit a negative temperature coefficient and in some cases to pass through a shallow minimum before achieving a constant value.

Two ordering transitions occur in Mn_3ZnC.* The knee in the ρ vs. T curve near 200° K (see Fig. 1) suggests an approach to constant resistivity (~320 micro-ohm cm) accompanying the lower transition. Complete spin disorder is not achieved at θ_N (~240° K), however, and the increase in resistivity is resumed as the temperature is raised through this point. Well above the upper transition temperature, the resistivity finally becomes constant at about 410 micro-ohm cm, but not before it has passed through a shallow minimum near θ_C. It is tempting to attribute this minimum to the short-range order effects referred to earlier.† This possibility is given added weight by the fact that the specific heat anomaly observed by Swanson at θ_C actually exhibits a very pronounced "tail" for $T > \theta_C$, usually a clear indication of the persistence of short-range order. Interestingly enough, the specific heat anomaly at θ_N has only a small tail. This anomaly, however, is larger overall than the upper one. While it has not been possible to separate completely these anomalies from one another or from the lattice heat capacity, it is evident qualitatively that the gain in both magnetic enthalpy and entropy, i.e. spin disorder, up to the lower transition point θ_N are larger than the corresponding quantities between θ_N and θ_C. This is consistent with the observation that between 0° K and 240° K the resistivity increases by a larger amount than it does between 240° K and 450° K and further supports the assumption that spin disorder provides the dominant scattering mechanism in this substance.

Measurements of the electrical resistivity of Mn_3ZnC between 4.2° K and 630° K are thus apparently interpretable on the basis of the spin-disorder scattering picture. This suggests that the electrons whose spins are primarily responsible for the magnetization are localized on manganese ions and that another group of electrons whose number is independent of temperature carries the electric current. This appears to be true both above and below the spin rearrangement transition temperature θ_N. There is no indication in the present results that delocalization of magnetic electrons occurs at θ_N as perhaps happens in magnetite at 120° K (Verwey, Haayman, and Romeijn 1947).

*Dysprosium also exhibits two magnetic transitions with corresponding resistivity anomalies (Coles 1958).

†Another example of such a minimum is that found for alpha-manganese at its Néel point (Coles 1958; White and Woods 1957).

REFERENCES

Brockhouse, B. N. and Myers, H. P. 1957. Can. J. Phys. **35**, 313.
Butters, R. G. and Myers, H. P. 1955. Phil. Mag. **46**, 132.
Coles, B. R. 1958. Advances in Phys. **7**, 40.
DeGennes, P. G. and Friedel, J. 1958. J. Phys. Chem. Solids, **4**, 71.
Swanson, M. L. To be published.
Verwey, E. J., Haayman, P. W., and Romeijn, F. C. 1947. J. Chem. Phys. **15**, 181.
White, G. K. and Woods, S. B. 1957. Can. J. Phys. **35**, 346.

A UNIFYING THEORY OF HIGH-LATITUDE GEOPHYSICAL PHENOMENA AND GEOMAGNETIC STORMS[1]

W. I. Axford and C. O. Hines

ABSTRACT

This paper is concerned with the occurrence at high latitudes of a large number of geophysical phenomena, including geomagnetic agitation and bay disturbances, aurorae, and various irregular distributions of ionospheric electrons. It shows that these may all be related in a simple way to a single causal agency, namely, a certain convection system in the outer portion of the earth's magnetosphere. The source of this convection is taken to be a viscous-like interaction between the magnetosphere and an assumed solar wind, though other sources of an equivalent nature may also be available. The model is capable of accounting for many aspects of the phenomena concerned, including the morphology of auroral forms and the occurrence of 'spiral' patterns in the loci of maximum intensities of several features. It also bears directly on the steady state of the magnetosphere, and in particular on the production of trapped particles in the outer Van Allen belt. In short, it provides a new basis on which a full understanding of these several phenomena may in time be built.

I. INTRODUCTION

Throughout the whole of the earth's exosphere, and through much of the underlying ionospheric region, the geomagnetic field exerts a strong control on the motion of ionized matter. This composite region of geomagnetic dominance is conveniently termed the magnetosphere, following Gold (1959).

It is common to think of the greater part of the magnetosphere in static terms, as a relatively calm expanse of quiescent plasma. But, as Gold has emphasized, there is no real reason for adopting such a view in advance and indeed convective motions of the ionization can be established without difficulty. Convection of the whole magnetosphere is in fact implicit in a variety of circumstances (including, for example, those described by the tidal dynamo theory) in which driving forces are exerted on the ionization at relatively low levels. The complementary case of convection generated at extreme altitudes has not been paid much attention as yet, but its consequences may be important in many phenomena.

We examine here one particular pattern of convection due to distant forces, whose natural occurrence we are led to expect. The driving mechanism that we envisage consists of a viscous-like interaction between the outer magnetospheric material and the interplanetary gas beyond. A similar convection could be established by other processes, and one such is proposed independently in an accompanying paper by Fejer (1961). Regardless of its manner of generation, however, the convective system we discuss has consequences of far-reaching import, and it is these that we wish to emphasize in the present paper.

[1]Manuscript received July 10, 1961.
Contribution from the Defence Research Telecommunications Establishment and the D.R.B. Theoretical Studies Group, Defence Research Board, Ottawa, Canada. The work was performed under projects PCC D48-95-11-01 and PCC D48-95-10-27.

The convection system will be relatively weak at times of low solar activity, but even then it provides a single comprehensive model to explain the high-latitude patterns of occurrence of magnetic agitation, sporadic-*E* and spread-*F* ionization, radio absorption, and radio-star scintillation centers. With a strengthened convection at times of high solar activity, the model can account for many features of auroral excitation, morphology, and motions, and of related magnetic variations. It provides, in fact, a new basis on which theories of detail in all these subjects may in time be built.

II. GENERAL DEVELOPMENT

II.1. The Basic Model

The outer boundary of the magnetosphere is in all probability determined in some way by an interaction between the geomagnetic field and the interplanetary gas that lies beyond. We have as yet no certain information on the density and motion of the interplanetary gas, however, so our picture of the outer magnetosphere must be tentative at best. For purposes of the present discussion, we shall adopt a model in which the interplanetary gas is generally in motion outward from the sun in a 'solar wind' (Parker 1958*a*), although we shall assume that the strength of this wind varies between quiet and disturbed conditions. We shall ignore the motion of the earth in its orbit, although we note at this point that such a motion could duplicate the essential feature (the viscous-like interaction) that we wish to abstract from the solar-wind model, even if the model should prove to be untenable in quiet conditions.

We must be concerned briefly with the distribution of geomagnetic field lines within the magnetosphere, especially in the case of the 'high-latitude' field lines that traverse the polar ionosphere. It is not known yet whether these field lines extend indefinitely into the interplanetary gas or whether they form closed loops from one polar region to the other, although the distinction is in some respects fundamental. The model we adopt here is one in which the high-latitude field lines do indeed form closed loops. These loops do not lie symmetrically about the geomagnetic axis, but instead are confined to a 'geomagnetic tail' that is formed in the 'lee' of the earth by the viscous effect of the assumed solar wind. This model has been described in its simplest form by Johnson (1960), while Piddington (1960*a*) has discussed a similar model for periods of geomagnetic storms and Beard (1960) has studied some details of its formation. Its topology is implicit in the model developed by Chapman and Ferraro for the initial phase of geomagnetic storms (Chapman and Ferraro 1931) and explicit in Dungey's model of rotating magnetic stars (Dungey 1958). It is sketched here in meridional section (Fig. 1) and in equatorial section (Fig. 2) following schematically the form depicted by Johnson. (The geomagnetic and geographic axes will be assumed coincident and perpendicular to the ecliptic plane for simplicity.)

It will be noted that, on this model, there is a bifurcation of field lines on the noon meridian: low-latitude lines are of a roughly dipole shape whereas high-latitude lines sweep round over the poles into the geomagnetic tail. Indeed, in discussions of the whole magnetic field it is convenient to think of

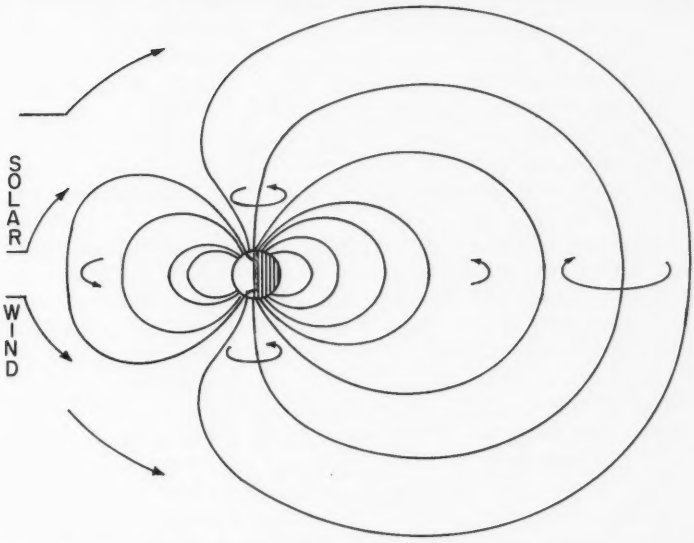


FIG. 1. A section of the magnetosphere, taken on the noon geomagnetic meridian. The solar wind (which is directed away from the sun) moves from left to right in the diagram, causing high-latitude field lines to be swept into the lee of the earth to form a geomagnetic 'tail'. The direction of rotation of the magnetosphere is indicated.

two classes of field lines, the one lying in a low-latitude torus and having a dipole-like shape, the other lying in the high-latitude torus that constitutes the geomagnetic tail. Field lines that rise on the day side of the earth near the boundary that separates these tori will be markedly deformed, and their allocation to one or other torus will by no means be clear-cut, while on the night side the distinction between the two tori essentially disappears. A region of confusion may be defined by the grossly distorted field lines that pass near the outer limits of the geomagnetic field, and this region maps along the field lines into two roughly semicircular 'zones of confusion' on the earth's surface, centered on the noon meridian. These zones are depicted in the accompanying figures.

In company with many others, and as already noted, we envisage a solar wind that varies in strength with solar activity. As the wind increases, more and more field lines are likely to be carried into the geomagnetic tail, and the zone of confusion would then proceed to lower latitudes. This behavior is reminiscent of the behavior of the auroral zone at times of high activity, and there is very likely a physical relationship between the two. Nevertheless, we resist the tendency to associate the zone of confusion with the auroral zone itself, and instead place it at somewhat higher latitudes. For example, our subsequent illustrations will be drawn on the basis of a transition from a low-latitude dipole field to a high-latitude magnetic tail at a latitude of 70°, some 3° or so above the conventional auroral zone. This transition corresponds

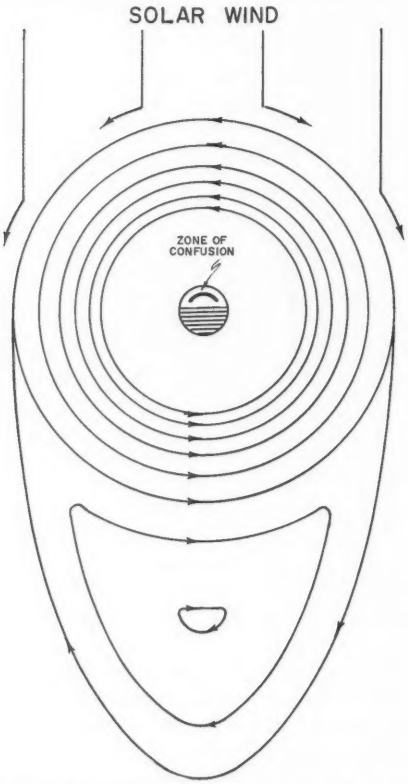


FIG. 2. An equatorial section of the magnetosphere viewed from above the north pole, showing streamlines for the case in which the magnetosphere is simply rotating in the manner depicted in Fig. 1. The streamlines are also equipotentials of the electric polarization that is induced by the rotation. Neighboring equipotentials are separated by 2 kilovolts.

to a cutoff in the magnetic field at a geocentric distance of 8 earth radii on the sunward side, using an unperturbed dipole extrapolation, or to a somewhat closer cutoff in a more self-consistent model. It is intended to represent moderately disturbed conditions, produced by a solar wind of perhaps 10 proton–electron pairs per cm^3 moving at a speed of 1000 km/sec.

II.2. The Impressed Convection

We now assume that the interaction between the outer magnetosphere and the solar wind contains a viscous-like component. By this we mean simply that some of the momentum of the solar wind is transferred across the boundary of the magnetosphere to the ionization within. The nature of this momentum transfer is, for present purposes, of minor importance; its existence, or the existence of an equivalent mechanism, is crucial.

We ourselves tend to think in terms of a coupling initiated by instabilities

at the boundary of the magnetosphere and carried to depth by a form of hydromagnetic eddy viscosity. The existence of the requisite instabilities has been anticipated theoretically by Dungey (1955; 1958, pp. 151–152) and Parker (1958*b*), while Sonett (1960) has discussed observational evidence of the subsequent coupling they produce. Sonett infers an appreciable exchange of momentum throughout the geocentric range 7–14 earth radii during relatively quiet conditions. Our subsequent illustrations indicate turbulence over the range $6\frac{1}{2}$–8 earth radii in the low-latitude torus for moderately disturbed conditions.

It is worth noting that other processes might produce the convective motions we deduce here, but that our later discussion is essentially independent of the manner in which the convection is established. The alternative processes include, for example, the charge separation deduced by Chapman and Ferraro (1933), the hydromagnetic interactions outlined by Dungey (1961), and the asymmetric injection of energetic charged particles now proposed by Fejer (1961). In fact, any mechanism that produces the *Ds* pattern of magnetic variations as a consequence of electromotive forces generated above the ionospheric *E* region must produce a pattern of magnetospheric circulation of the form we shall be discussing here.

In the presence of a viscous interaction, which we now assume for purposes of discussion, the solar wind in its passage past the earth will carry material in the outer regions of the magnetosphere away from the sun, towards the geomagnetic tail. This material, being linked to the geomagnetic field, is constrained to remain within the magnetosphere and so it builds up an excess concentration in the region of the tail. The build-up cannot persist in a steady state, and consequently a return flow of ionization must be set up in the interior of the magnetosphere. Circulatory loops will be established, of a form depicted in Fig. 3, in which the ionization convects continuously. In the popular terminology of hydromagnetics, the ionization will be 'frozen' to the geomagnetic field lines (see Dungey 1958, for example) and will carry them along in the convective motion. The field lines can be distorted only slightly in the process, except perhaps at the outer boundaries, and hence the whole of each field line will carry out the convective motion imposed upon it. The ionization at lower heights, being frozen onto the field lines, will similarly convect. Hence the pattern of motion whose equatorial section is shown in Fig. 3 will be extended throughout the body of the magnetosphere along the field lines. It is this gigantic convective system that forms the central feature of our subsequent discussion.

The foregoing (hydromagnetic) description may not satisfy all tastes, and we shall rephrase it accordingly in more classical terms. We begin again with the basic assumption that the ionization of the outer magnetosphere is set into motion in the direction of flow of the solar wind, and that it establishes a certain steady-state velocity pattern (**V**). This motion of the electrons and ions takes place in the presence of the geomagnetic induction (**B**), however, and it can therefore proceed only if an electric polarization field (**E**) is established by the accumulation of space charge to offset the Lorentz force. (In

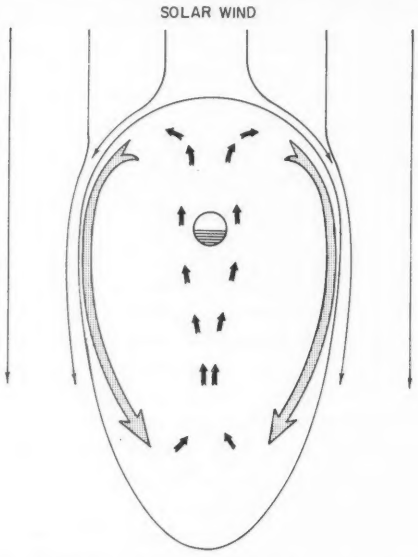


FIG. 3. The motion impressed on the magnetosphere by a viscous-like interaction with the solar wind. This is an equatorial section of the magnetosphere with the solar wind blowing from top to bottom. The viscous-like interaction causes tubes of force which lie near the surface of the magnetosphere to be pulled around into the geomagnetic tail, as indicated by the large arrows. A return flow (indicated by the small arrows) takes place in the interior of the magnetosphere.

the models of Chapman and Ferraro (1933) and of Fejer (1961) the process of charge separation is more fundamental, but it would induce in turn the motions now under discussion.) The polarization field must be derivable from a potential (ϕ) in the steady state, whence

$$\text{grad } \phi = -\mathbf{E} = \mathbf{V}\times\mathbf{B}. \tag{1}$$

(Rationalized m.k.s.a. units are used throughout the formal development.) This relation limits the type of velocity field that can be established in the magnetosphere, for it implies that $\mathbf{V}$ must everywhere lie in the local equipotential surface, and that that surface must be so oriented as to include the local $\mathbf{B}$ vector.

Consider now a section cut through the magnetosphere in a surface orthogonal to $\mathbf{B}$, in the (magnetic) equatorial plane, for example. The equipotential surfaces will cut this surface in some pattern of closed curves which are themselves equipotentials. Since the $\mathbf{V}$ vectors lie in the equipotential surfaces, their components ($\mathbf{V}_\perp$) orthogonal to $\mathbf{B}$ must lie along the equipotential curves and so must trace out flow lines that are closed convective loops. The simplest and most readily justifiable pattern of such loops consistent with the impressed motion is of the form presented in Fig. 3, which we therefore recover by the present reasoning. This pattern will be examined in more detail shortly.

Since equation (1) implies that the geomagnetic flux lines are equipotentials, it also implies that the electric fields existing in one region must be mapped along the flux lines to other regions. It then implies that the magnetospheric ionization in those regions is also convecting. Indeed, it is not difficult to show from equation (1) and Maxwell's relations that all the ionization that lies along a tube of magnetic flux at one time must convect together, so as to remain always along some common tube of flux. This is the classical analogue of (part of) the 'frozen field' concept of hydromagnetics, and it recovers the earlier conclusion that the motion of ionization depicted in Fig. 3 will be mapped throughout the magnetosphere along magnetic flux lines.

Certain factors tend to inhibit the magnetospheric motions, and these warrant brief comment. In the first place, energy would be required or given up when a change is made from one steady state to another. This energy would have to be transferred across the magnetospheric boundary, and presumably is so transferred when, for example, a strengthening solar wind blows more and more field lines into the geomagnetic tail. But in the steady state, the convective motions that are of interest here contain far less energy than does the geomagnetic field. Accordingly, any field deformation that is produced by the convection itself should be negligible (except perhaps in the outermost levels, as already noted, where the convective energy density might be comparable with the geomagnetic energy density). In hydromagnetic parlance, this conclusion implies that a tube of magnetic flux must vary its cross section as it convects, in order to maintain the strength of the field at its appropriate equilibrium value. The 'frozen in' ionization must similarly alter its cross section, and so a two-dimensional divergence of $\mathbf{V}_\perp$ will be produced. (This latter statement may be verified on classical grounds by the application of equation (1) and Maxwell's relations once again.) This introduces a strong tendency for the magnetospheric matter to vary in density and pressure as it convects, despite the possibility of ionization 'leaking out' into the lower ionosphere, and such variations raise the question of stability.

The foregoing points have been stressed by Gold (1959) in an extensive discussion of the principles of magnetospheric convection. On the matter of stability, he notes that certain regions (defined by shells of magnetic field lines) might permit rapid adiabatic convection, while others, unless strongly forced, might permit only slow isothermal convection; still others might be highly stable against any convective processes, and the Van Allen belts may be particularly important in this regard. He suggests, however, that the outer magnetosphere is likely to be of such a nature as to permit the more rapid convection, and that this convection might affect certain geophysical features—the very ones, in fact, that will come under close study in our subsequent development. We assume here that stability will not be a major problem in the outer regions that are of primary concern in the present study.

Atmospheric viscosity and ohmic losses become appreciable near the base of the ionosphere and act to inhibit the motion. Being purely resistive in nature, they cannot prevent the convection but can only limit its speed. We must assume that the observed velocities are determined by the balancing of

energy transfer at the magnetospheric boundary with energy loss in these dissipative processes, once any inherent stability has been overcome. This assumption will bear further analysis at another time, but two preliminary checks on it can be made quite readily. The losses are mainly ohmic, and amount to something like 10^{-7} watts/m^3 for typical auroral currents. This loss extends throughout a height range of the order 10^4 m over an area of the order 10^{14} m^2, for a total dissipation rate of the order 10^{11} watts. The local rate of energy dissipation is down by an order of magnitude from the rate introduced by normal photo-ionization processes, which is comforting, while the total rate is down by an order of magnitude or more from the rate of energy capture implied by Sonett's (1960) calculations. The uncertainties inherent both in the present estimates and in Sonett's leave open the possibility that an even closer agreement could be obtained to imply that the greater part of the energy incident from the solar wind is in fact dissipated by ionospheric currents with magnetospheric convection acting as an intermediary. Such a conclusion would overcome the difficulty encountered by Sonett in the search for a sink for the energy influx he deduced.

At this point, we might add some credence to our views on a viscous-like interaction by the following comments. We shall later argue that the motion of auroral irregularities corresponds to the motion of associated geomagnetic field lines (using the hydromagnetic terminology). The observed speeds are normally somewhat less than 1 km/sec, and these would be scaled up to something like 20 km/sec for the speed of the corresponding field lines at their greatest height above the earth. This speed is still substantially below the speed associated with the solar wind at disturbed times, as it must be if the operating mechanism is indeed viscous-like.

Turning now to the details of the convection pattern, we note that these will depend in large measure on the stability and the dissipative processes discussed above. For example, the return flow within the body of the magnetosphere would tend to be channelled through regions of low rather than high stability. Energy dissipation at ionospheric heights results from current flow, and the currents must be driven by forces in the magnetosphere. These forces would include not only the basic interaction with the solar wind, but secondary mechanisms such as those provided by ion and electron pressure gradients. The net steady-state flow would be determined by the interplay of these several features, and would be governed by a formidable array of relations of which equation (1) is only a single member and an approximate one at that.

We shall make no attempt here at any analytical approach to this difficult problem. Instead, we shall simply sketch a refined version of Fig. 3 in a form which seems reasonable to us, which is based in large part on observational evidence (assuming our interpretation of the observational material to be true), and which is compatible with the physical principles of the problem. The proposed pattern is presented in Fig. 4. Certain of its features warrant discussion.

Our subsequent comparison with observational data leads us to place the innermost part of the return flow at a geomagnetic shell whose position (as

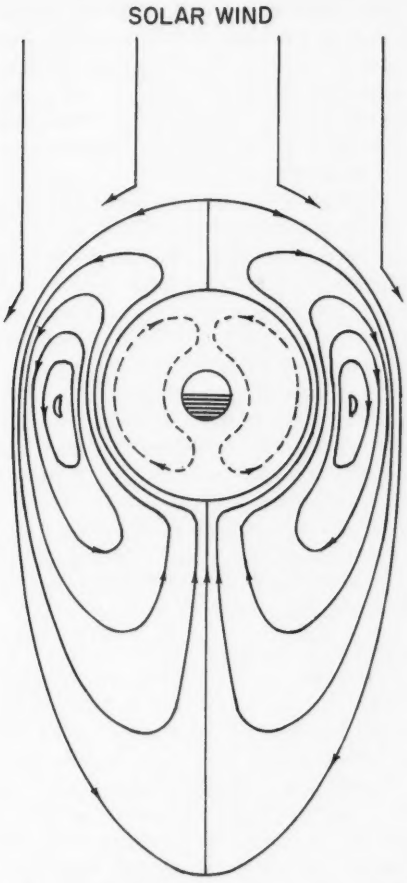


FIG. 4. The proposed pattern of streamlines (or alternatively of the equipotentials of the electric field, at separations of 2 kilovolts) in the equatorial plane of the magnetosphere. The main circulation does not penetrate closer than about $4\frac{1}{2}$ earth radii; dotted lines within this region indicate a possible inner convection system.

measured at the base of the field lines) is somewhat below the conventional auroral zone, at a latitude of about 62°. This corresponds to a return flow bounded on its inner side at a geocentric distance of $4\frac{1}{2}$ earth radii in the equatorial plane. These values are probably representative of moderately disturbed conditions and of some suitable mean between lower numbers at times of strong disturbance and higher numbers during relatively quiet periods.

We would associate the inner boundary region with some form of stability, and we consider it most likely that the energetic particles of the outer Van

Allen belt provide the stabilizing agency. The suggestion that this belt is found at lower latitudes on the night-side of the earth (Rees and Reid 1959) would imply in turn that the circulatory system could penetrate to greater depths and lower latitudes at night. A similar conclusion might be drawn for other potential stabilizing processes, and it would help to explain the morphology of aurorae. It is, however, a refinement that we shall not attempt to incorporate in our illustrations.

We may also note in passing that the circulatory system will to some extent be self-stabilizing, for it results in the energization of particles during the inward convection as will be discussed. The energization tends to increase the longitudinal drift motions that are introduced by the inhomogeneity of the geomagnetic field, and these motions then tend to carry the convecting particles around in longitude out of the zone of inward convection.

In addition to the main convective patterns already discussed and depicted by solid contours in Fig. 4, we note the possible existence of an inner circulatory system represented by broken contours. This inner circulation complicates the suggested pattern in a manner analogous to certain simply hydrodynamic systems, and it might be set up by some viscous (hydromagnetic) interaction at the inner boundary of the outer circulatory system or by some other secondary process such as will be described subsequently. It is postulated here primarily for the explanation of the low-latitude system of *Ds* currents, although it may have other useful implications as well. It will be discussed only briefly in the subsequent development.

II.3. Superimposed Effects of Rotation

The discussion to this point has treated the convection as a departure from a static condition. In fact, of course, the earth and its ionosphere are in rotation, and the magnetospheric ionization will also rotate (as a consequence of hydromagnetic coupling) provided drag effects imposed by the interplanetary gas are not too great. The rotation of the high-latitude torus takes place about its own axis, rather than the earth's, as Johnson (1960) has described, and is freed from the inhibiting effects that would have arisen (Hines 1959) had the high-latitude field lines been taken to extend into the interplanetary gas. On the present model, then, it seems reasonable to assume that rotation proceeds without significant impediment except perhaps at the very boundaries of the magnetosphere where the viscous interaction is imposed.

As with the convective motion, the rotational motion can proceed only in the presence of an electrostatic field. The equipotential surfaces of the field will again lie along the magnetic flux lines and the rotational flow lines. Their equatorial sections have been depicted already in Fig. 2.

There must be some interaction between the rotational motion and the convective flow, but its evaluation lies beyond the scope of the present study. We shall be content with a derivation of the net flow produced simply by superimposing the two independent velocity fields. The superpositioning can be achieved most simply by treating the contours of Figs. 2 and 4 as equipotentials, assigning values of potential to individual curves in each set of

contours, adding the two values of potential at a number of points throughout the field, and drawing new contours of constant total potential. The assignment of values to the original contours is straightforward in the case of the rotational system, at least for the undistorted portions of the geomagnetic field, since the angular velocity and the field are known. The corresponding assignment for the counter-rotating geomagnetic tail and for the convective system is rather arbitrary in the absence of a detailed analysis, although the direction of increasing potential is predetermined by the direction of **V** required. Despite this arbitrariness, the general trend of the resultant flow pattern is quite unambiguous; it is depicted for one specific case in Fig. 5. The principal feature to be noted is that the ionization convecting inward on the nighttime

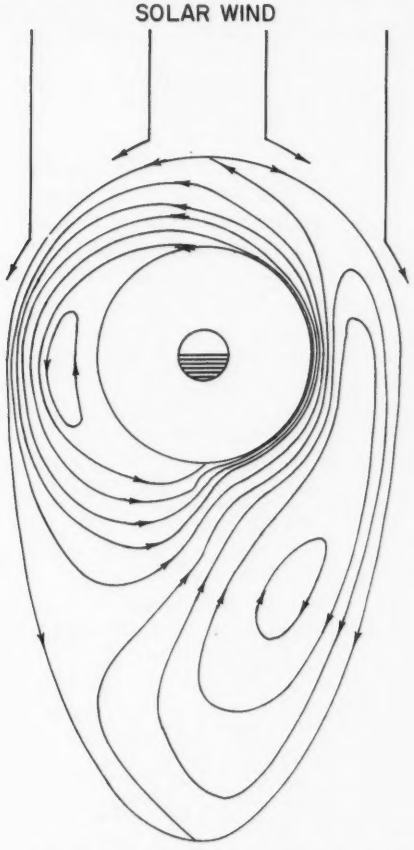


FIG. 5. A composite picture obtained by superimposing the equipotentials of Figs. (2) and (3), showing the type of motion to be expected in the presence of rotation and a viscous-like interaction between the solar wind and the surface of the magnetosphere.

side of the earth is carried by the rotation around to and through the morning side. Other features will be pointed out later, as they become pertinent to the discussion.

II.4. The Motion of Energetic Magnetospheric Ionization

The convective motion described so far has taken no account of the gyrations of individual particles about the geomagnetic field lines or their random motions along those lines. Such motions lead to additional drift components, as is now well known in connection with so-called trapped radiation. The additional drift will not be important for particles of sufficiently low energy, but it becomes quite comparable with the convective motions for energies of the order of a few kiloelectron volts. Particles of greater energy still will move predominantly with their natural drift velocities, and will be subject only to small perturbations from their natural orbits as a consequence of the

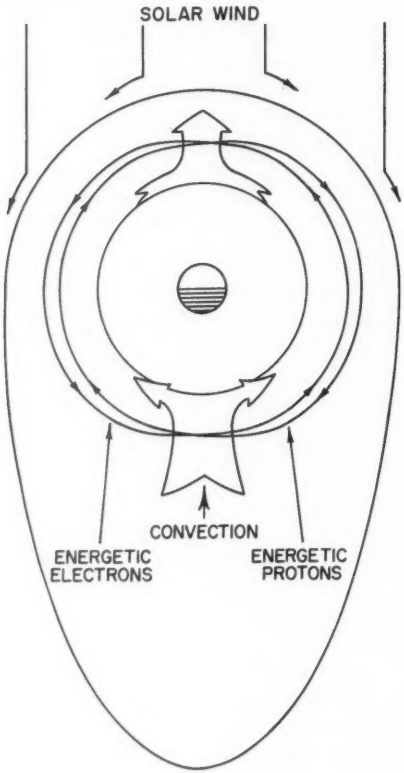


FIG. 6. The effect of the proposed circulation pattern on energetic trapped particles. A thin shell of such particles is split into separate electron and proton shells, such that the electrons drift on lower-latitude field lines during the morning hours and the protons during the afternoon hours.

magnetospheric convection (or, more properly, as a consequence of the electric fields that accompany that convection). Since proton and electron drifts proceed in opposite senses around the earth in longitude, while the convective effects act in the same sense at any one point, a given shell of (mono-) energetic protons and electrons will be split into two deformed shells, one of protons and one of electrons, as depicted in Fig. 6.

It is clear that this effect will tend to produce a charge separation of the energetic particles, even when integrated over a range of energies and through a continuous distribution of shells. Fejer (1961) points out that this separation will modify the convective system, and that it will do so in such a way as to produce the inner loops we tentatively drew in Fig. 4. That figure should now be considered as a composite, including the primary convective system and any superimposed secondary effects of the charge separation. The latter tend to be only transitory, however, since in a steady state the separation of energetic charged particles would be largely offset by the accumulation of low-energy space charge; some residual motions, having a complicated form, would remain.

II.5. The Induced Ionospheric Motions

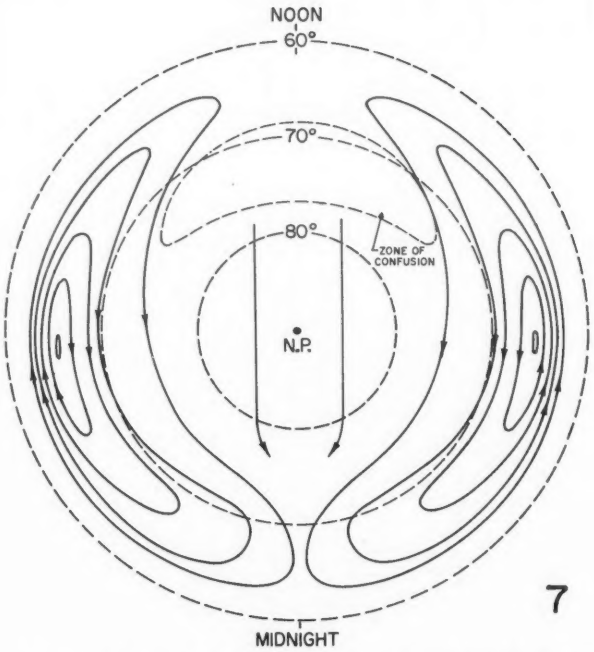
The convection systems that have so far been depicted by equatorial sections take on a somewhat different appearance when extrapolated down the geomagnetic field lines. The equipotential surfaces have sections at ionospheric heights that are best illustrated in a polar view of the earth, as in Fig. 7 for the convection alone and in Fig. 8 with rotation superimposed. The outermost regions of the magnetosphere map into the 'zone of confusion' previously introduced, and this zone is shown explicitly. The flow pattern in and near the zone is difficult to deduce because of the severe distortion of the pertinent field lines. We have contented ourselves in the present representation with a mapping of the flow pattern in the more certain regions at a distance from this zone, and with an implied continuity of flow across it. Figures 7 and 8 show the equipotentials accordingly. Residual deviations from this pattern may have some bearing on geophysical observations, however, and should be kept in mind.

In the ionospheric F region, equation (1) remains valid to a high degree of accuracy, and it continues to imply a convection of ionization along the equipotential surfaces. The same is true for electrons in the E region, down to heights of (say) 100 km. In the case of ions, however, collisions become important and must be taken into account. There results a deviation of the ion motion from the simple equipotential contours.

The effect is perhaps best seen from the equation

$$0 = e(\mathbf{E}+\mathbf{V}_i\times\mathbf{B}) + M_iK_i(\mathbf{V}_n-\mathbf{V}_i), \tag{2}$$

which is a better approximation than (1) to the equation of motion of ions. (Here e is the charge of an ion, M_i is its mass, and $\mathbf{V}_i$ its velocity; $\mathbf{V}_n$ is the velocity of the neutral gas, and K_i is a 'frictional frequency' related to the ion–neutral collisional frequency ν_{in} by $M_n\nu_{in} = (M_i+M_n)K_i$ where M_n is the mass of a neutral molecule. Inertial and certain other terms are still missing



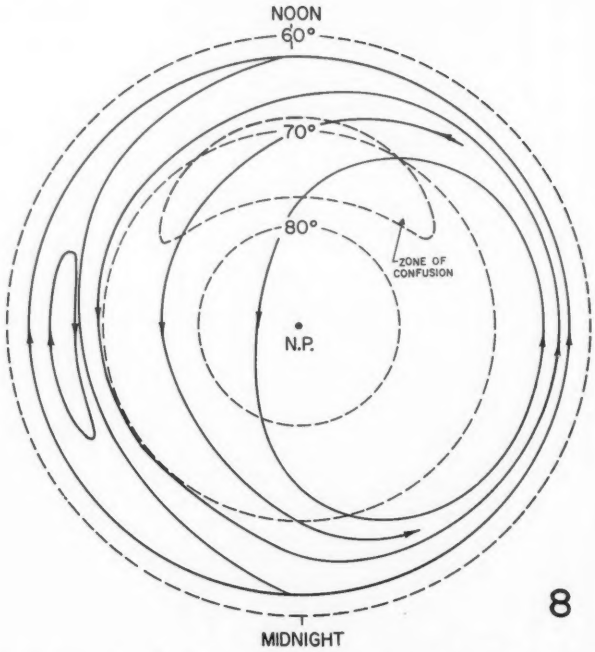


FIG. 7. The pattern of motion at ionospheric levels, obtained by mapping the streamlines of Fig. 4 down onto the northern hemisphere along the lines of force of the geomagnetic field.

FIG. 8. The pattern of motion at ionospheric levels, obtained by mapping the streamline of Fig. 5 down onto the northern hemisphere along the lines of force of the geomagnetic field.

from the equation given.) We may think first of the case when $V_n = 0$, and then note that (1) is likely to be a good approximation to (2) only if $eB/M_iK_i \gg K_i$, that is, if the ion gyro-frequency greatly exceeds the ion frictional frequency. The two become comparable in the upper part of the E region, however, and in the lower part of this region the gross inequality is in fact reversed. In the latter circumstances, $\mathbf{V}_i$ approximates to $e\mathbf{E}/M_iK_i$ and is considerably reduced in magnitude from the value E/B attained by ions at higher levels and by electrons even in the E region. The diminished velocity relative to electrons implies that the electrons will become the main current-carrying constituent, and that the current will then flow roughly along the equipotentials but in the sense opposite to the motions indicated in Fig. 7. (Figure 7 is appropriate in this connection rather than Fig. 8, because the rotational component of $\mathbf{V}_i$ in the collisional term of equation (2) is offset by an equal component in $\mathbf{V}_n$ there, and (2) may be replaced once again by (1) insofar as the rotation is concerned.)

These conclusions require more exact analysis in terms of the anisotropic conductivity tensor that applies at E-region heights, and with the inclusion of resultant polarization fields. The net effect would be a height-integrated effect, and its nature would depend on height-integrated conductivities. The latter have been deduced by various authors (for example, Baker and Martyn 1953) for application to atmospheric-dynamo theories, and it has been found that the Hall conductivity predominates except at equatorial latitudes. This conclusion leads, in effect, to the description just given in which the electrons are the dominant carriers of current.

The development of polarization fields in certain circumstances will lead to a flow of current in regions not directly affected by the convection process, and this secondary current can in turn generate secondary convection systems. This complication provides another possible source of the internal circulation depicted in Fig. 4, and of the associated low-latitude *Ds* current system.

Unlike atmospheric-dynamo theories, in which the motion of the neutral atmosphere is prescribed, the present theory involves the neutral gas with convective motions that are impressed on it by ions. These may be of significance under certain conditions. They are most simply studied by application of the approximate equation of motion

$$N_nM_n d\mathbf{V}_n/dt = N_iM_iK_i(\mathbf{V}_i - \mathbf{V}_n) \tag{3}$$

in which the driving force is taken to be solely the reaction on the neutral gas of the collisional term in (2); N_n and N_i are the number densities of neutral particles and ions respectively. The implications of this equation are in turn most simply studied under the assumption that $\mathbf{V}_n$ and $\mathbf{V}_i$ vary as exp $(i\omega t)$, and that $d/dt \simeq \partial/\partial t$; then

$$\mathbf{V}_n = \mathbf{V}_i(1 + i\omega\tau_K)^{-1} \tag{4}$$

where

$$\tau_K \equiv N_nM_n/N_iM_iK_i. \tag{5}$$

It is evident from this that the convective motion of the neutral gas is likely

to be negligible only for variations on a time scale much shorter than τ_K, which is of the order of some tens of minutes in the F region and some hours in the E region.

For variations with a time scale longer than τ_K on the other hand, it is evident that the neutral molecules can be carried along by the ions in their motion, provided no other forces interfere. The residual difference of velocities, for $\omega\tau_K \ll 1$, is given by

$$\mathbf{V}_n - \mathbf{V}_1 \simeq -i\omega\tau_K \mathbf{V}_1 \tag{6}$$

from (4), and this result combines with (2) to yield

$$0 \simeq e(\mathbf{E}+\mathbf{V}_1\times\mathbf{B}) - i\omega\tau_K M_1 K_1 \mathbf{V}_1 \tag{7}$$

as the approximate equation of motion for the ions. We now see that the ions will follow the convective motion (1) provided $eB/M_1 \gg \omega\tau_K K_1$, which is a less stringent condition than the earlier requirement that $eB/M_1 \gg K_1$. Indeed, it is evident that the ions may convect closely with the electrons in motions whose time scale is at least as long as

$$\tau_B \equiv N_n M_n / N_1 eB, \tag{8}$$

even at heights where the ion collision frequency exceeds the ion gyrofrequency. The quantity τ_B varies throughout the E region and varies from day to night; it is of the order of a few hours at 125 km during the day, and it increases from this value at night and at lower levels unless abnormal sources of ionization are introduced. We are led to expect the duration of any intense current flow to be limited by this effect. Residual currents will flow for times exceeding τ_B, but they will be less intense and their pattern may differ appreciably from that of the equipotentials. These aspects will be discussed again in relation to the observational data.

II.6. The Distribution of Turbulence

As has been noted already, it is expected and observed that the outer portions of the magnetosphere are turbulent. Our present model indicates that the material in these regions will be convected to the interior, and we therefore expect a corresponding convection of the turbulence. The turbulence energy will of course decay as the convection proceeds, but this may be more than offset by the effects of the compression that accompanies transport to the interior. Such a compression would result in increased absolute fluctuations of ionization density and magnetic field, if the fractional fluctuations remained essentially unchanged, and in increased gradients due to the diminishing linear dimensions. The way is then open for the acceleration of particles to high energies (Parker 1958*b*, 1961), and for the production of severe hydromagnetic disturbances. It is clear that a variety of geophysical phenomena might follow, and the pattern of convected turbulence is therefore of some consequence.

This pattern is depicted in Fig. 9, which retains the major convective system of Fig. 5 and introduces the following regions for convenience of discussion:

I. The outer magnetospheric regions in which the turbulence is originally generated. Here the turbulence exists independently of the circulatory motions,

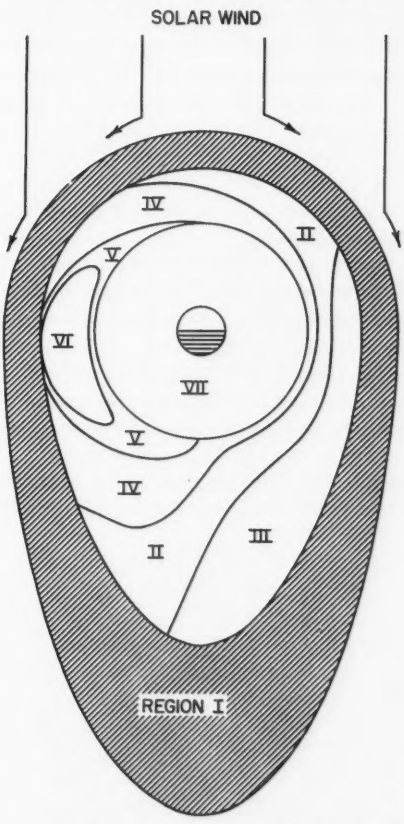


FIG. 9. The distribution of turbulence in the equatorial plane of the magnetosphere. The hatched area denotes the region in which there is intimate contact with the interplanetary gas and into which momentum is transferred directly by the solar wind. The various zones are defined in the text.

and it may be expected to decrease in intensity from the outer boundary inwards.

II. The regions through which the outermost material of I is convected. This material is likely to be the most turbulent initially, and it is subject to strong compression. It may then be expected to be of major concern.

III. The region between I and II on the morning side. The material here is convected from the interior of I, and is not severely compressed. It should therefore be turbulent, but not of the same importance as II.

IV and V. Regions containing turbulence convected from the interior of I on the evening side. IV loops round the earth on the morning side, and V on the evening side, although V will not exist unless the convection is sufficiently strong in relation to the rotation. The material of IV is not as turbulent as

that of II initially, but it is subject to a stronger compression; it therefore ranks with II in its possible geophysical consequences. The material in V is even less turbulent initially and is subject to a slightly less severe compression than that of IV, but the compression is still strong and its turbulence may therefore be of significance on those occasions when the region itself exists.

VI. A region that is essentially devoid of turbulence, since the material contained in it never convects through I. If the convection is sufficiently weak, so that V does not exist, VI extends from I to the inner portion of the magnetosphere on the evening side.

VII. The inner region into which the circulation does not penetrate. Turbulence generated on the surface of the magnetosphere is unimportant in this region, which corresponds to geomagnetic latitudes less than about 62° on the surface of the earth.

The regions depicted in Fig. 9 map along the field lines into specific zones at ionospheric heights and ground level. These are illustrated in Fig. 10, where the numbering system of Fig. 9 is retained. The 'zone of confusion' previously introduced lies centrally along zone I in this figure.

In view of a widespread interest in 'spiral' patterns in high-latitude geophysical phenomena, it is worth noting even at this stage the appearance of

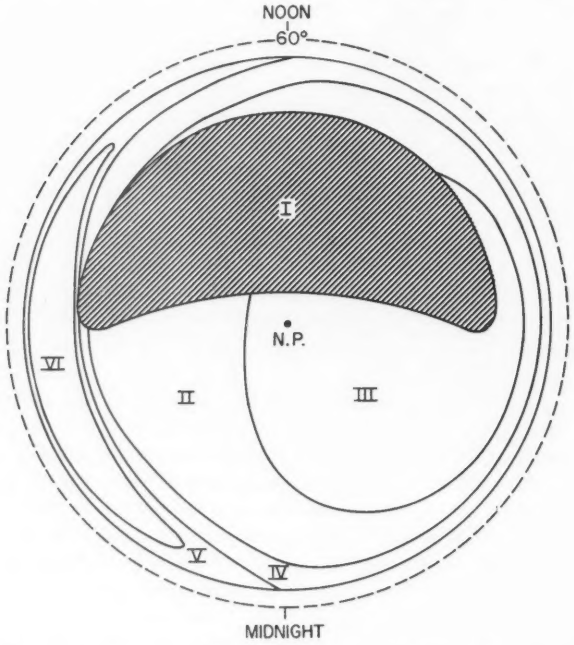


FIG. 10. The distribution of turbulence in the magnetosphere as observed at ionospheric levels. This pattern is obtained by mapping the distribution of Fig. 10 down along lines of force of the geomagnetic field.

spiral-like forms in zones II, IV, and V. This feature will receive more discussion directly, but it quite clearly has no relation to the 'Störmer spirals' that are so often invoked. The latter can have little direct bearing on the observed phenomena in any event (Agy 1960).

III. GEOPHYSICAL APPLICATIONS

III.1. Features Associated with Magnetospheric Turbulence

Whenever sufficiently intense turbulence occurs in the magnetosphere, its effects should be observable by means of magnetic and ionospheric measurements at ground level. In particular, we might expect such diverse phenomena as geomagnetic agitation, auroral activity (both visual and radio), auroral absorption, high-latitude types of sporadic-E and spread-F ionization, and high-latitude radio-star scintillations to be associated in part with magnetospheric turbulence. Indeed, the field-aligned irregularities that characterize several of these phenomena can perhaps best be explained as a consequence of just such a turbulence, mapped along the geomagnetic field lines by particle precipitation and hydromagnetic propagation. The details of this mapping process we postpone for separate investigation and turn instead to a comparison of the theoretical distributions and motions with readily available observational data.

In some cases, the distributional information has been summarized in the form of geographic loci of maximum intensity of the feature under consideration. While such summaries run the risk of being misleading, particularly when more than one maximum is found, they are nevertheless convenient and adequate for our immediate purpose. Summary data of this type are displayed in Fig. 11 for geomagnetic agitation (Burdo 1957; or see Hope 1961), intense high-latitude sporadic-E ionization (Hagg *et al.* 1959, December data), spread-F ionization (L. E. Petrie, private communication), visual aurorae (Malville 1959), 'auroral-type' radio scattering (Forsyth *et al.* 1960), and auroral absorption (L. E. Montbriand, private communication).

It is readily apparent from this display that a major enhancement of disturbance phenomena is to be found in a single broad swath that sweeps from higher to lower latitudes over a range of longitudes on the night side of the earth. This duplicates in form our combination of zones II and IV in their compression phase, and our earlier arbitrary (but justifiable) choice of boundaries for the magnetospheric convection was such as to ensure duplication of the latitude range as well. We are led to conclude that the dominant distribution of high-latitude disturbance phenomena can be explained adequately as a consequence of convected turbulence, following the pattern given by our model.

The two secondary spirals provided by the observational data in the morning and afternoon sectors are not so readily explained. The characteristics of magnetic agitation on these spirals are sufficiently different from those of the nighttime loop that some authors have sought a separate mechanism for their source (for example, Hope 1961). Comparison with Fig. 10 then leads us to suggest that they are associated with zone I, and hence with the primary

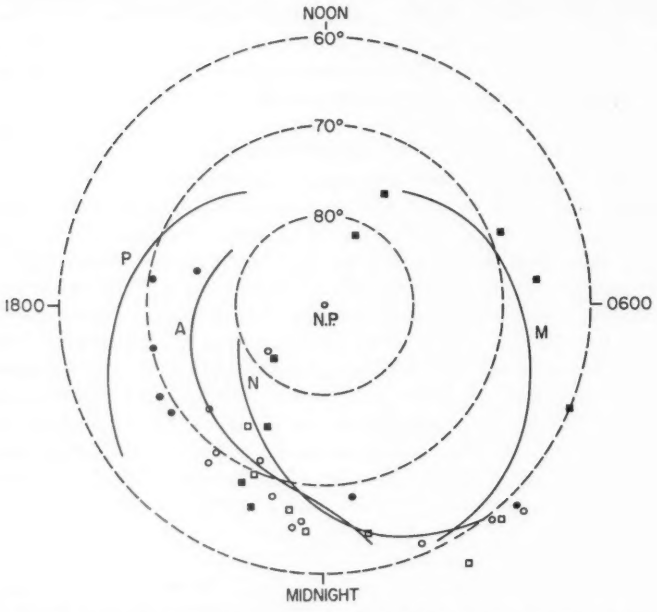


FIG. 11. The positions of diurnal maxima of various disturbance phenomena in geomagnetic co-ordinates. The loci P, N, and M refer to geomagnetic agitation and A to visual aurora; ○ to sporadic-E; □ to radio aurora; ● to spread-F; ■ to auroral radio absorption.

unconvected turbulence of the outer magnetosphere, perhaps combined to some extent with the recovery phases of zones II, IV, and V. This view cannot be strongly supported, however, prior to a fuller examination both of the theory and of the observations.

Scintillations of radio stars, due apparently to ionization irregularities high in the F region, have not been analyzed in the fashion of Fig. 11. But observations made at subauroral latitudes, related to ionization at low auroral latitudes, do reveal a peak in scintillation index near geomagnetic midnight (Hewish 1952; Dagg 1957). The index rises rapidly in the late evening hours, and falls gradually during the morning. This behavior appears to be consistent with the dominant role of the combined zone, II+IV. The earlier onset of scintillations at times of high activity (Dagg 1957) would follow naturally upon the expansion of the circulatory loops of Fig. 10 to lower latitudes, and upon the increased role that might be played by zone V.

Certain evidence supports the suggestion that there is a region of low disturbance, namely the non-turbulent zone VI, situated on the evening side of the auroral zone. For example, a representative mean diurnal variation of ionospheric absorption has been published for College, Alaska, at a geomagnetic latitude of 65° (Little and Leinbach 1958). While it does not depict any clear localized maxima, it does reveal a very marked minimum of absorption some

4 or 5 hours before geomagnetic midnight. The diurnal variation of sporadic-E at Narssarssuak (geomagnetic latitude 71°) shows a broad and pronounced minimum of occurrence at this time (Olesen and Wright 1961). The non-turbulent zone would of course also be associated with the quiet auroral arcs that characterize the pre-midnight period in contrast to the much more active forms of the early morning hours (for example, Davis 1960).

Our picture of convecting turbulence leads us to expect that any given magnetospheric irregularity would convect more or less with the pattern of flow depicted in Fig. 5. The corresponding ionospheric irregularities would then follow the pattern shown in Fig. 8, or, when measured relative to the rotating earth, that shown in Fig. 7. The available evidence supports this conclusion.

The evidence derives in part from measurements of radio-star scintillations. The ionization irregularities, which lie at low auroral latitudes in the cases of interest, move predominantly westward before geomagnetic midnight and eastward after (Hewish 1952; Briggs 1960). The tendency for a reversal at midnight becomes more pronounced as the magnetic activity increases, which suggests that the convective systems we have been discussing may be competing with some other mechanism (for example, the tidal dynamo) at the relatively low latitudes concerned. Other evidence to this effect can be adduced, but is of no real importance here.

Much more complete observations are available from auroral studies. Measurements made at low auroral latitudes reveal again the tendency for a westward motion before midnight and an eastward motion after, in both visible and radio aurorae (for example, Meinel and Schulte 1953; Kim and Currie 1958; Bullough and Kaiser 1955; Lyon and Kavadas 1958). It has also been established explicitly (Kim and Currie 1958) that no systematic westward motion is introduced by the eastward rotation of the earth, and this conclusion would be difficult to explain unless an eastward rotation were associated with the generating mechanism. Such a rotation is, of course, inherent in the model we have developed, and so the observational evidence is compatible with the implications of Fig. 7.

The pattern of motion of auroral irregularities has been established more fully by Davis (1960), in an analysis of IGY data. His work includes observations made at high auroral latitudes, where the east–west motions are found to reverse just as the motions depicted in Fig. 7 reverse. Indeed, the over-all pattern of motion presented by Davis is essentially duplicated by the pertinent portions of Fig. 7. Moreover, the alignment of auroral forms in the sun–earth direction at geomagnetic latitudes greater than about 80° would suggest that the flow lines for this region shown in Figs. 7 and 8 are essentially correct, if, as at lower latitudes, the auroral forms tend to lie along flow lines. (The explanation of this tendency lies beyond the scope of the present paper, as does the explanation of auroral forms themselves. We suspect that both follow from a form of instability in the associated currents, and this possibility is now being examined in more detail.) Davis' results have been recently confirmed by Denholm (1961) and Stoffregen (1961).

The description of the motions of aurora and related phenomena can be

completed by allowing for additional north–south motions. We have suggested, for example, that the outer radiation belt prevents the circulation from penetrating deeper than about $4\frac{1}{2}$ earth radii in the equatorial plane of the magnetosphere. Due to longitudinal variations in the strength of the geomagnetic field, the particles forming the belt tend to move to lower latitudes on the night side of the earth (Rees and Reid 1959). This effect can be expected to result in a north–south motion of the low-latitude boundary of auroral activity as a whole, the motion being towards the equator in the evening and away from the equator in the morning. Such motions have been described in practice by Bond (1960) for example. This movement, which may amount to several hundred kilometers, is best observed from stations outside the auroral zone; it should not be confused with the north–south motion of irregularities within an auroral form, which do not necessarily follow the same pattern.

Finally, we might note an obvious fact: the increase in the speed of both scintillation and auroral irregularities that accompanies increased magnetic activity (for example, Dagg 1957) is immediately explained in the present model, as due simply to a more rapid convection induced by a strengthened momentum interchange at the magnetospheric boundary. The latter could be produced by an increase in either the strength or the irregularity of the solar wind.

In summarizing all these points, it may be said that convected turbulence is clearly capable of accounting for the distribution and motion of a wide variety of high-latitude geophysical features. We believe that such an explanation is likely to prove correct.

III.2. Features Associated with the Large-Scale Convection and Compression

In addition to their apparent enhancement of turbulent features, magnetospheric convection and compression are capable of more direct effects on a larger scale. In the case of convection, these can come about by the displacement of energetic particles already trapped in the geomagnetic field (as described in II.4) and by the transport of energetic solar particles from the boundary of the magnetosphere to its interior. The compression, which is in fact an inherent aspect of convection in an inhomogeneous field, corresponds to an energization of both types of particle.

The energization process derives from a motion of ions and electrons across the equipotential surfaces of ϕ, and it therefore implies a violation of equation (1). The cause of this violation is to be found in the drift motion that ionization naturally attains in the inhomogeneous geomagnetic field, and this motion is proportional to the energy of the ionization amongst other things. Whether the drift speed is large or small in comparison with the convective speed then depends on the energy: low-energy ionization tends to follow the flow lines of the convective system, while high-energy ionization tends to follow its natural drift trajectories. A transitional energy (W_t) may be envisaged at which the two components of motion are comparable in magnitude. Its value varies with the strength of the convective system and with geocentric distance, but it lies typically in the range 1–100 kev.

For the lower-energy ionization in particular, it is often useful to consider the energization process as one of adiabatic compression (Gold 1959). If losses through precipitation are ignored, ionization convected from a geocentric distance R_1 to a distance R_2 will be subject to compression by a factor $(R_1/R_2)^3$ through the diminution of its cross section normal to the (dipole) geomagnetic field, and by a further factor of the order of (R_1/R_2) due to the diminution of the length of a field line above the precipitation level. Precipitation into the lower ionosphere will reduce these factors, but they provide a preliminary basis for discussion. If raised to the power $\gamma-1$ (where γ is the adiabatic index, or specific heat ratio of thermodynamics) they yield the appropriate factors for the energization process.

If interparticle collisions are important, the effective γ is 5/3 and the energy factor becomes $(R_1/R_2)^{8/3}$. If collisions are unimportant, γ is 2 for the transverse compression and 3 for the longitudinal compression (Spitzer 1956), so the energy of transverse thermal motions will be altered by the factor $(R_1/R_2)^3$ and that of the longitudinal motions by $(R_1/R_2)^2$. The convected low-energy ionization evidently alters its geocentric distance by a factor of 3 or so in the course of its circulation, so its energy can be increased by an order of magnitude or more.

The higher-energy ionization cannot attain such a large proportionate enhancement of its energy, for the maximum ratio of R_1/R_2 it can attain is limited by its natural drift velocity as discussed immediately above. In this case, it is possible to put a limit on the absolute amount of energy that a given particle can acquire, in the collision-free case, since the energy derives essentially from the electrostatic field **E**. This maximum energy (W_m) will be of the order $eVBL$, where L is a representative length scale and the other parameters have been defined already. At auroral heights $B \sim 10^{-4.5}$ weber/m^2, whence $W_m \sim 10^{-4.5} V_a L_a$ electron volts, where V_a is typically $10^{2.5}$ m/s and L_a may be taken as 10^6 m roughly, and so $W_m \sim 10$ kev. This value might easily be increased or decreased by an order of magnitude, depending on the strength and extent of the circulation. It may be no coincidence that $W_m \sim W_t$, although the significance of this conclusion has yet to be assessed.

The particles of the outer magnetosphere may be expected to exhibit a broad range of energies, from a fraction of an electron volt to solar-wind proton values of several kev. The solar-wind protons are expected to share their energy with electrons in a relatively short time (T. Gold, private communication), thus providing a supply of electrons having energies of the order of 1 kev. The convection will act on all particles to carry them inwards, although the more energetic will of course break free as they tend towards the energy W_t. The inwardly convecting stream will then act as a continuous source of particles of about this energy, throughout the night half of the magnetosphere. Electrons of this energy are capable of penetrating to auroral heights, and they may therefore be looked to as a potential source of aurorae and of abnormal E-region ionization.

The precipitation of energetic particles in fact depends upon a reduction of their mirror heights, and this takes place as the particles are carried inwards

on the night side of the earth. The strength of the magnetic induction (B_M) at the mirror point of a particle whose velocity components at the equator are $v_\perp$ and $v_\parallel$, perpendicular and parallel to the field respectively, is given by

$$B_M = B_E(v_\perp^2+v_\parallel^2)/v_\perp^2, \tag{9}$$

where B_E is the strength of the magnetic induction at the equator on the field line of interest. If collisions are negligible, the ratio $B_E/v_\perp^2$ remains constant during compression. The resultant B_M is increased in proportion to the increasing energy of the particle and at the same time the field line on which this B_M must be attained is moving to lower latitudes. Both effects lead to a depression of the mirror height, which in total may be quite considerable.

The foregoing effects are of course more significant for the softer component of the particle distribution (particle energies $\lesssim$20 kev, say) because of the limit W_m placed on the acquisition of energy, and we look to this component as the source of most auroral forms. The harder component will not be unaffected, however, and we expect that it is important in certain stable emissions. In particular, we suggest that the very stable band of H_α emissions that accompanies aurorae is due to the precipitation of trapped protons of moderately high energies ($\sim$100 kev, say) subject to the effects just described. The H_α emissions are to be found at latitudes below the lowest-latitude auroral forms in the pre-midnight period, and then move to greater latitudes where they normally become lost in the general auroral emission (L. E. Montbriand, private communication). This crossover is in the sense to be expected from a mapping of the proton loop in Fig. 6 down to auroral heights.

It is perhaps worth noting that the convection and compression discussed here tend to favor the precipitation of protons in the late evening and electrons in the early morning. This is a consequence of the fact that energetic protons drift westward, and energetic electrons eastward, from the region of inward convection centered on midnight (cf. Fig. 6). There is observational evidence of just such a bias in auroral emissions (Rees *et al.* 1961).

The precipitation of both convected and trapped particles would of course be enhanced by the action of turbulence, with implications that have already been examined. But the mechanisms discussed in the present section are not dependent on turbulence, and so can proceed even in the non-turbulent region VI of Fig. 9. They can then be responsible for the quiet auroral forms and glow that occur in zone VI of Fig. 10 (and elsewhere), and we believe them to be a major contributing factor at least.

The fact that all our mechanisms—turbulence, convection, and compression —tend to favor the production of aurorae at night is possibly too obvious to warrant comment, although it is an achievement not often attained by auroral theories in the past.

We do not propose to develop the fine points of our model at this stage, although it is easy to see that the theory allows considerable scope for the understanding of more detailed features than we have yet discussed. We would emphasize that our mechanisms of auroral production are by no means exclusive, and we would expect a final description of auroral phenomena to include many aspects of theories other than (but compatible with) our own.

III.3. Magnetic Disturbances

The large-scale magnetic disturbances that contribute to geomagnetic storm variations are often thought of in four parts: the initial-phase world-wide increase, the main-phase world-wide decrease, the diurnally varying *Ds* component, and short-lived 'bays'. Some writers consider the *Ds* variation to represent simply a statistical averaging of the bay disturbances, and its separate identity certainly has not been clearly established.

There is widespread agreement that the initial phase of a storm is a direct consequence of solar particles approaching the geomagnetic field, in the manner studied first by Chapman and Ferraro (1931) and more recently recast in hydromagnetic terms (for example, Dessler and Parker 1959). There is also agreement that the main phase could be caused by the enhanced circulation of trapped energetic particles (Singer 1957; Dessler and Parker 1959; Akasofu 1960), in the form of a 'diamagnetic ring current', although the means of providing an enhancement have been by no means clear. Piddington (1960*a*) in particular has challenged the processes so far advanced for this purpose.

A direct application of the development in the preceding section indicates a new process, and a very powerful one at that. The compression of magnetospheric ionization, and its consequent energization, would lead to an enhanced circulation of the type required. This enhancement would be offset to some extent by the decompression and de-energization that occurs on the day side, and by the loss of magnetospheric ionization through enhanced precipitation; whether the balance would be favorable or not is a question that must be bypassed for the present. But, in addition to the normal magnetospheric ionization, material captured from the solar wind will be caught up in the convection as previously described, and its contribution must be in the sense required for the main phase. We therefore suggest that the convective system provides the mechanism of introducing solar particles to the interior of the magnetosphere, and that of supplying thereby the storm-time ring current. This mechanism avoids the difficulties discussed by Piddington, whether these are real or not.

In this picture, we would interpret the delayed onset of the main phase (by a few hours, following the start of the initial phase) as the time taken before substantial solar material is convected to the interior and energized. The subsequent decay of the main phase could follow from the capture process discussed by Dessler and Parker (1959), which is independent of the initial generation process.

Turning now to the high-latitude *Ds* system and bay disturbances, we recall the *E*-region current systems that would be established by convection as discussed in Section II.5. As a first approximation, these currents will flow along contours of constant potential in the sense opposite to the 'frozen' convective motions of electrons. Chapman's analysis (1956) of the *Ds* system provides a pattern of current flow which is essentially equivalent to the convective pattern of Fig. 7, and with the direction of flow reversed as required.

During storms of greater intensity than those analyzed by Chapman, it appears that the diurnally varying current systems are shifted to somewhat earlier hours and intensified at auroral latitudes, strongly after midnight and

weakly after noon (Silsbee and Vestine 1942). These intensified currents are more often associated with 'bays' than with the *Ds* system, but the semantical schism is perhaps best avoided, when possible, until the phenomenon itself is understood.

We advance two suggestions in explanation of the change, although we recognize that the problem is extremely involved in detail and no simple solution is likely to be complete in itself. In the first place, we note that the flow of strong currents is dependent on the neutral atmosphere remaining free from the convective motion of the ionization, as discussed in II.5, and that this freedom may be lost after a period of a few hours (τ_K) if the convective forces are sufficiently strong. As the neutral atmosphere of the *E* region rotates through the midnight position at auroral latitudes, it enters a new circulatory loop to which its motion is not adjusted, and strong currents can flow. By dawn, however, it can have been accelerated to take part in the new convection, and the current strengths fall. The transitional sequence is repeated in the hours immediately following noon, when relatively strong currents flow once again.

The second process applies in particular to the post-midnight enhancement and may account for its relatively great strength. It involves simply an enhanced conductivity due to increased ionization, resulting from the combined effects of the turbulent zone II+IV and the compression and convection as discussed previously.

In both cases, and indeed in the case of the quiet *Ds* system, a divergence of ionization flow must be anticipated. This will result in a secondary system of electric fields, and in a secondary current flow, such that the *Ds* system of Fig. 7 has superimposed on it dipole-like current systems for which the direction of the dipole is that of the undistorted local *Ds* current. The nature of these secondary currents has been discussed by Martyn (1951), who has shown that the low-latitude *Ds* system is likely to be produced as a result. While we do not agree with Martyn's method of establishing the initial current flow, as will be discussed, we do consider the secondary electric fields to be important and to be capable of producing the low-latitude current system. The secondary fields would in turn modify the pattern of convection we have been discussing so far, and in particular they could establish a convection of the inner magnetosphere of the type indicated in Fig. 4.

On the basis of this theory the *Ds* current system should be regarded as being present as long as there is a solar wind and simply becoming greatly enhanced during geomagnetic storms. However, a number of effects may cause the pattern to become (partly) apparent even during relatively quiet periods, such as the enhancement of conductivity as a result of sudden ionospheric disturbances.

As an aside, it may be noted that Gold (1959) has suggested that a low-latitude convection is established normally, as a consequence of strong equatorial heating. This would be expected to produce a rising motion in daytime, outward through the magnetosphere, and it was suggested as a source of low-latitude ionization irregularities. Without examining the merits

of this view, we would simply remark that such a convection would be opposed and inhibited by the low-latitude *Ds* convection just discussed, and that the occurrence of low-latitude spread-*F* ionization is indeed inhibited at times of geomagnetic disturbance (Lyon *et al.* 1960).

Returning now to auroral latitudes, we may note that the intensity and sense of bay disturbances are related to the speed and direction of motion of auroral irregularities (both visual and radio; for example, Heppner 1954, Harang and Tröim 1960). In particular they exhibit positive or negative changes in the horizontal geomagnetic field according as the auroral motion is westward or eastward. This association represents a correlation in detail of the over-all statistical patterns already discussed, and it is confirmed in even more detail by the work of Kaiser (1958) on extremely short-lived bays. It is exactly what would be anticipated on the simple basis we have adopted, in which convection at *E*-region heights is largely confined to the electrons alone. In the case of the short-lived bays, we would probably have to associate the enhanced speed of motion with turbulence, but the same result would follow nevertheless.

As with the geophysical features previously examined, we will not attempt to develop our model in detail at present nor to incorporate in it many attractive aspects of other studies. Again it is evident, however, that the model does include properties whose further examination and amplification may help to elucidate the details observed.

III.4. The Equilibrium State of the Magnetosphere

We have suggested that, even at quiet times, a circulation of the type we have discussed will be present. If so, it should be of significance to the equilibrium state of the magnetosphere. Two examples will serve to illustrate the role it may play.

First, we consider the ionization at thermal energies that constitutes by far the largest part of the material of the magnetosphere. It has been thought in the past that the distribution of this ionization above the *F* region is governed by the hydrostatic law, suitably modified to take account of ambipolar and rotational effects (for example, Johnson and Fish 1960; Öpik and Singer 1960). However, the circulatory motion is likely to be quite comparable in magnitude with the thermal motion of ions, and is therefore likely to affect the distribution appreciably.

In an extreme case, in which ion flow into and out of the *F* region was prevented, the ion content of a tube of geomagnetic flux would remain unaltered as it convected, and a steady equilibrium could be established only if that content were proportional to the flux contained in the tube. Near the equatorial plane this would result in a density roughly proportional to R^{-4}, where R is geocentric distance; at other latitudes the distribution would be of a more complicated functional form. This differs substantially from the hydrostatic result, and is therefore incompatible with it. We would expect the actual distribution to be determined by an appropriate combination of the pertinent hydrostatic and convective forces.

We turn now to the more energetic ionization that constitutes the radiation belts, and consider in particular the 'soft' component of the outer belt in which particle energies are of the order of 20 kev. We have already noted that energies of this order can be produced by the convection at times of auroral activity, and even a substantially weaker circulation could be of consequence. If the circulation proceeded slowly and quietly, these particles would be de-energized in the course of the subsequent outward convection, and they would make no net contribution to the trapped component. In the presence of turbulence, however, the pertinent adiabatic invariants break down, and a substantial portion of the energized particles can be further energized and trapped as shown by Parker (1961).

Parker stresses the point that turbulence at the outer boundary of the magnetosphere is not sufficient, for more energized particles would be diffused outward and lost than would be introduced and retained in the process. On the other hand, a strong turbulence in the interior, surrounded by a relatively quiet region, would be ideal for the production of trapped particles. If compressed turbulence is as effective as our previous discussion implied, it seems to provide precisely the type of situation required for the effective operation of Parker's process. Moreover, it provides optimum conditions close to the inner boundary of the main convective system, at a geocentric distance of 4–5 earth radii which is essentially the position of the maximum flux in the soft component of the outer belt. We are led to suggest that this component is due to particles of solar origin, convected to the interior and there trapped by the effects of turbulence. We envisage a strong injection during severely disturbed conditions, but also a lower level of injection at quiet times such as will maintain the equilibrium flux density against losses to the atmosphere.

The variations that are found at times of magnetic storms also receive a fairly natural explanation on the present basis. The flux of detected particles first shows a marked decrease and recovers after some hours to levels substantially above normal (Arnoldy *et al.* 1960). The preliminary decrease has been associated by some with the general compression of the geomagnetic field during the initial phase of the storm, but we should now note that the precipitation produced by an enhanced convection, and an associated increase of W_m, would provide an alternative mechanism, and one which might well be dominant. (If it were, we would expect the spectrum of particle energies to harden at this time, since the convection process acts more strongly on the soft component; but this conclusion might also follow from the suggested general compression.) In this regard, it is important to note that the enhancement of the convection should proceed almost simultaneously throughout the magnetosphere, being delayed in the interior only by the short propagation time of hydromagnetic effects, so the related enhancement of precipitation would begin with the onset of the storm.

On the other hand, particles captured from the strengthened solar wind would reach the interior only after the relatively long convection time of a few hours, and would only then provide an enhanced rate of injection into the particle belts. This rate of injection would reach a maximum when the available solar particles and the turbulence reach a maximum, presumably at about

the time of the maximum in the main phase. We would then expect the integrated contribution to the detectable flux to reach its maximum at some later time still, when the enhanced injection rate returns once again to its equilibrium value. The anticipated sequence of events is illustrated by Fig.12.

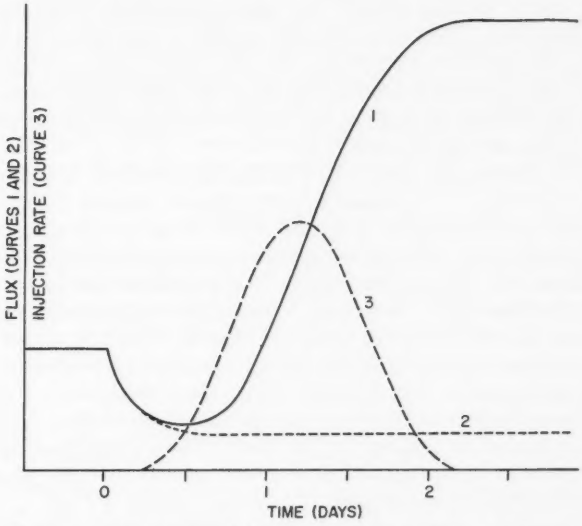


FIG. 12. The temporal variation of the counting rate in the outer Van Allen belt (1). The curve (2) represents the counting rate of particles which formed the belt at the time of the sudden commencement, according to the description of events given in the text, and (3) represents the injection rate required to produce (1). The variation of (3) is essentially that of the storm-time ring current.

IV. CONCLUDING DISCUSSION

IV.1. Comparison with Related Theories

The model we have presented appears to be the first in which magnetospheric convection plays an important role. Indeed, except for Gold's emphasis on it and its potential, this convection has been totally overlooked in the past. It is nevertheless implicit in any model that drives the *Ds* current system from above. Martyn's (1951) theory is a prominent representative of such models, and a brief examination of it in contrast to our own can be instructive.

Martyn employs the charge separation of Chapman and Ferraro as his driving mechanism, and we have already indicated that it could serve equally for ours. The great extension of the earth's exosphere was not appreciated at the time Martyn wrote, so the convection of ionization within it played no part in his development. Convection in the *F* and *E* regions did, however, and Martyn deduced a pattern of circulation similar to our Fig. 7. But in his derivation the sense of flow was opposite to ours, and this forced him to invoke a greater mobility for positive than for negative ions in order to account for the sense of the *Ds* current system. Such an assumption is unjustified theoretically and is opposed by the correlation of bay disturbances with auroral

motions and by the over-all trend of pre-midnight and post-midnight motions actually observed.

The discrepancy between Martyn's theory and ours can be traced to the fact that Martyn assumed the charge-separation process to occur at a geocentric distance of only 5 earth radii, and to drive the circulation at auroral latitudes directly. In our model, the viscous-like driving agency acts at distances greater than 6 or 7 earth radii, and operates directly only on the regions linked to zone I in Fig. 10. The motion across the rest of the polar cap and at auroral latitudes then constitutes a return flow, opposite in direction to the flow determined by Martyn. As we have seen, this new pattern is fully compatible with the observations.

The theory developed recently by Piddington (1960*a*, *b*) is closely related to ours, in that it adopts a viscous-like interaction to produce the geomagnetic tail (and indeed, was the first to do so) and to drive the *Ds* current system. In both cases, however, Piddington appears to employ the viscous interaction only to deform the field and exospheric ionization from one static condition to another. He therefore fails to exploit, or even to recognize, the important consequences that convection can produce. In this regard, it should be noted that any distinction between a distorted static geomagnetic field and a distorted convecting geomagnetic field is purely a semantical distinction, although the second of these descriptions conjures up a picture of events that is more likely to be useful if properly applied. In the case of the ionization, of course, the distinction between a static and convective system is perfectly real, and we are forced to believe in convection.

In this connection, we should comment on Piddington's view of the main phase of a geomagnetic storm. If we understand him correctly, Piddington pictures this as a consequence of an expansion of the low-latitude torus of the geomagnetic field at a time of diminishing solar-wind pressure, with the expansion acting to fill a gap left by field lines that were swept into the geomagnetic tail. It would seem that the enhanced viscous interactions must be maintained at the same time, in order to keep the field lines in the tail, or that some stability of the tail must be assumed. While we agree that this process might play some part, we would note that the necessary assumptions cannot be strongly defended and that their failure would certainly limit the importance of the mechanism. But, what is more important, we note that convection through the interior can return the field lines of the tail to their initial configuration in the low-latitude torus, without running counter to the viscous drag of the solar wind (and indeed being aided by it). Accordingly, we have serious doubts that this process could compete effectively against the ring current discussed in our preceding section. As the latter is free from the difficulties that have been associated with ring currents in the past, we consider it likely to be the most dominant mechanism.

We need not pursue this comparison of theories further, except to note again that mechanisms capable of driving the *Ds* current system from high in the magnetosphere should lead to convective motions of the type we have discussed. Theories of auroral and geomagnetic storms that do not include

such mechanisms may also imply convection, but the circulatory patterns would likely be of a different form or orientation; their comparison with the present model lies beyond our immediate scope.

IV.2. Summary

We have discussed here a mechanism for the generation of a convective motion in the outer magnetosphere and have examined very broadly some of its implications. The latter are in good agreement with, and provide a forthright explanation of, a wide variety of features in high-latitude geophysical phenomena and storm-time variations. We are led to believe that the convection system is of major importance to these phenomena, and we expect it to provide a new basis on which theories of detail may in time be based.

ACKNOWLEDGMENTS

Discussions with many of our colleagues have been of great help in the formulation of the foregoing theory. We would like to acknowledge in particular the valuable comments of Dr. J. A. Fejer at all stages in the development, and of Prof. T. Gold and Dr. G. C. Reid in specific areas.

REFERENCES

AGY, V. 1960. J. Atmospheric and Terrest. Phys. **19**, 136.
AKASOFU, S.-I. 1960. J. Geophys. Research, **65**, 535.
ARNOLDY, R. L., HOFFMAN, R. A., and WINCKLER, J. R. 1960. J. Geophys. Research, **65**, 1361.
BAKER, W. G. and MARTYN, D. F. 1953. Phil. Trans. Roy. Soc. A, **246**, 281.
BEARD, D. B. 1960. J. Geophys. Research, **65**, 3559.
BOND, F. R. 1960. Australian J. Phys. **13**, 477.
BRIGGS, B. H. 1960. U.R.S.I. XII General Assembly.
BULLOUGH, K. and KAISER, T. R. 1955. J. Atmospheric and Terrest. Phys. **6**, 198.
BURDO, O. A. 1955. Defence Research Board Translation by E. R. Hope T321R (1959).
CHAPMAN, S. 1956. Vistas in Astron. **2**, 912.
CHAPMAN, S. and FERRARO, V. C. A. 1931. Terrestrial Magnetism and Atmospheric Elec. **36**, 171.
——— 1933. Terrestrial Magnetism and Atmospheric Elec. **38**, 79.
DAGG, M. 1957. J. Atmospheric and Terrest. Phys. **10**, 194.
DAVIS, T. N. 1960. J. Geophys. Research, **65**, 3497.
DENHOLM, J. V. 1961. J. Geophys. Research, **66**, 2105.
DESSLER, A. J. and PARKER, E. N. 1959. J. Geophys. Research, **64**, 2239.
DUNGEY, J. W. 1955. Proc. Ionosphere Conf. (Phys. Soc. London), 229.
——— 1958. Cosmic electrodynamics (Cambridge University Press).
——— 1961. Phys. Rev. Letters, **6**, 47.
FEJER, J. A. 1961. Can. J. Phys. **39**. This issue
FORSYTH, P. A., GREEN, F., and MAH, W. 1960. Can. J. Phys. **38**, 770.
GOLD, T. 1959. J. Geophys. Research, **64**, 1219.
HAGG, E. L., MULDREW, D., and WARREN, E. 1959. J. Atmospheric and Terrest. Phys. **14**, 345.
HARANG, L. and TRÖIM, J. 1960. Planetary and Space Sci. **5**, 33.
HEPPNER, J. P. 1954. J. Geophys. Research, **59**, 329.
HEWISH, A. 1952. Proc. Roy. Soc. A, **216**, 494.
HINES, C. O. 1959. J. Geophys. Research, **65**, 141.
HOPE, E. R. 1961. J. Geophys. Research, **66**, 747.
JOHNSON, F. S. 1960. J. Geophys. Research, **65**, 3049.
JOHNSON, F. S. and FISH, R. A. 1960. Astrophys. J. **131**, 502.
KAISER, T. R. 1958. Ann. géophys. **14**, 76.
KIM, J. A. and CURRIE, B. W. 1960. Can. J. Phys. **38**, 1366.
LITTLE, C. G. and LEINBACH, H. 1958. Proc. I.R.E. **46**, 334.
LYON, G. F. and KAVADAS, A. 1958. Can. J. Phys. **36**, 1661.
LYON, A. J., SKINNER, N. J., and WRIGHT, R. W. H. 1960. J. Atmospheric and Terrest. Phys. **19**, 145.

MALVILLE, J. M. 1959. J. Geophys. Research, **64**, 1389.
MARTYN, D. F. 1951. Nature, **167**, 92.
MEINEL, A. B. and SCHULTE, D. H. 1953. Astrophys. J. **117**, 454.
OLESEN, J. K. and WRIGHT, J. W. 1961. J. Geophys. Research, **66**, 1127.
ÖPIK, F. J. and SINGER, S. F. 1960. Phys. Fluids, **2**, 653.
PARKER, E. N. 1958*a*. Phys. Fluids, **1**, 171.
——— 1958*b*. Astrophys. J. **128**, 664.
——— 1961. J. Geophys. Research, **66**, 693.
PIDDINGTON, J. H. 1960*a*. J. Geophys. Research, **65**, 93.
——— 1960*b*. Geophys. J. **3**, 314.
REES, M. H. and REID, G. C. 1959. Nature, **184**, 539.
REES, M. H., ROMICK, G. J., and BELON, A. 1961. Planetary and Space Sci. **5**, 87.
SILSBEE, A. H. and VESTINE, E. H. 1942. Terrestrial Magnetism and Atmospheric Elec. **47**, 195.
SINGER, S. F. 1957. Trans. Am. Geophys. Union, **38**, 175.
SONETT, C. P. 1960. Phys. Rev. Letters, **5**, 46.
SONETT, C. P., JUDGE, D. L., SIMS, A. R., and KELSO, J. M. 1960. J. Geophys. Research, **65**, 55.
SPITZER, L. 1956. Physics of fully ionized gases (Interscience Publishers, New York).
STOFFREGEN, W. 1961. J. Atmospheric and Terrest. Phys. **21**, 257.

STUDY OF THE EFFECT OF GAMMA RADIATION ON THE MOLECULE OF PENTOBARBITAL[1]

J. A. R. Cloutier, Bruce C. Flann, and Jane M. Manson

ABSTRACT

Modern instrumental techniques were used to investigate the effect of gamma radiation on the molecule of pentobarbital. The results were correlated with each other and with those from a biological test. The main effect of radiation appeared to be centered about the two substituents of the molecule at the number 5 position.

INTRODUCTION

A preliminary experiment had shown that a gas was produced when pentobarbital was irradiated in the dry state by Co^{60} gammas. Identification of the many components of this gas corroborated the radiation damage mechanism suggested by an electron paramagnetic resonance study of pentobarbital prepared and irradiated in a boric acid glass (Cloutier 1961). It is the purpose of this report to present the experimental findings and correlate them with other data from chemical and biological tests on gamma-irradiated pentobarbital sodium.

Irradiation of the Samples

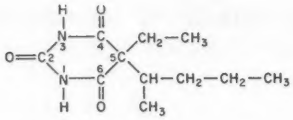
The samples were irradiated in a 500-curie Co^{60} Transfer Case Irradiator obtained from the Commercial Products Division of the Atomic Energy of Canada Limited. The average dose rate was 220,000 rads/hour. Absolute dose-yield determinations were not attempted in this experiment. Samples were simply irradiated at a level which was known experimentally to yield concentrations which allowed study of the given effect. Exposure times varied from a few hours to as long as $12\frac{1}{2}$ months.

For the gas analysis experiments, pentobarbital was irradiated in a sealed pyrex container which had been either evacuated or filled with nitrogen at atmospheric pressure. The pyrex containers were designed to permit sampling of the gaseous phase with a minimum of losses or contamination.

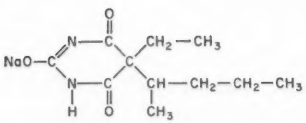
Material

The free acid and sodium salt of pentobarbital were commercial products of high purity. Levi and Hubley (1956) used samples from the same lot number to characterize pentobarbital and study its copper-pyridine complex. The 4,6-di-*p*-nitrobenzyl (Chatten and Levi 1957; Cloutier and Manson 1959), the 1,3-dimethyl (Manson and Cloutier 1961), and the 4,6-dixanthyl (Flann and Cloutier 1961) derivatives of pentobarbital were prepared and characterized by members of this laboratory. Specimens from the same preparations were available for this study.

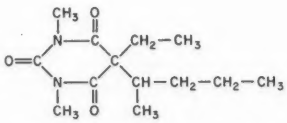
[1]Manuscript received June 26, 1961.
Contribution from the Food and Drug Laboratories, Department of National Health and Welfare, Ottawa, Canada.



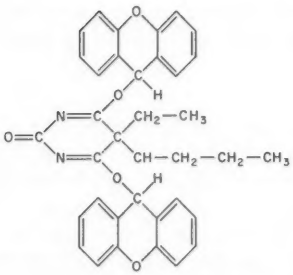
nembutal ®
pentobarbital (U.S.P.)
pentobarbitone (B.P.)
5-ethyl-5-(1-methylbutyl) barbituric acid



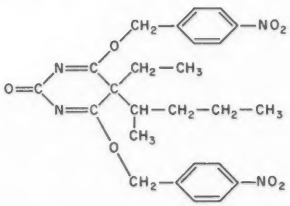
pentobarbital sodium



1,3-dimethyl pentobarbital



4,6-dixanthyl pentobarbital



4,6-di-p-nitrobenzyl pentobarbital

FIG. 1. Structural formulas of pentobarbital and some of its derivatives.

The structural formulae of the above compounds are given in Fig. 1. It will be noted that the structural formula for the dixanthyl derivative differs from that suggested by McCutcheon and Plein (1949). A recent infrared study of the reaction of barbiturates with xanthydrol (Flann and Cloutier 1961) indicated that the two xanthyl groups were attached to positions 4,6 of the malonylurea ring and not to positions 1,3 as previously assumed.

INFRARED DATA

The gas produced during irradiation of pentobarbital, exposed $12\frac{1}{2}$ months to Co^{60} gammas, was analyzed with a Perkin–Elmer Model 21 infrared spectrophotometer equipped with sodium chloride optics. The gas sample was introduced into an evacuated 7.5-cm infrared cell having sodium chloride windows and the gas pressure inside the cell was adjusted to atmospheric pressure with dry nitrogen. The spectrogram of Fig. 2 was obtained in 25 minutes at a

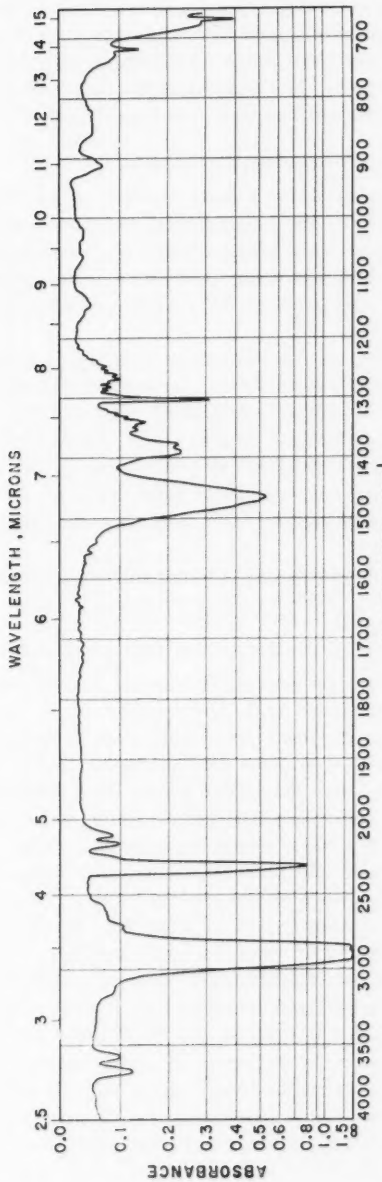


FIG. 2. Infrared spectrogram of the gas produced during irradiation of pentobarbital by Co^{60} gammas.

resolution setting of 927. The results reported in Table I were based on correlations with Sadtler standard gas spectra, the spectra of Pierson *et al.* (1956), and the spectra from gases obtained from the Matheson Company.

Two other experiments were made with pentobarbital sodium irradiated for 3 weeks and $1\frac{1}{2}$ months. The same gaseous components were identified; however, the relative proportions were not exactly the same.

GAS CHROMATOGRAPHIC DATA

Measured volumes of the same radiation-produced gas were also analyzed with a Perkin–Elmer Model 154-C gas chromatograph. Two types of Perkin–Elmer partition columns were used: column A, diisodecyl phthalate and column J, silica gel. The operating conditions were generally as follows: helium carrier at 10-lb inlet pressure; 25° C; 2-meter column; detector voltage, 8 volts.

The results pertaining to the $12\frac{1}{2}$-month irradiation of pentobarbital are presented in Table I. For identification purposes, the observed elution times were compared with those of gases generated (CO_2 and H_2) or obtained from either the Matheson Company (CO and CH_4) or the Phillips "66" Hydrocarbons Company (ethane, propane, *n*- and iso-butane, *n*- and iso-pentane). The estimates of relative concentrations were made by comparing the peak areas of the sample chromatograms with those of standard ones obtained with known volumes of gases.

The same gaseous components were found following the 3-week irradiation of pentobarbital sodium.

ELECTRON PARAMAGNETIC RESONANCE DATA

The effect of radiation on pentobarbital and five of its derivatives was studied at room temperature with the method of electron paramagnetic resonance. The spectrometer used has already been described (Cloutier 1961). Irradiation of pentobarbital in air or in a nitrogen atmosphere, Figs. 3C and 3D, did not give rise to any detectable signal. Pentobarbital sodium, when irradiated in air or in a nitrogen atmosphere, Figs. 3G and 3H, exhibited a weak singlet, 20 oersteds wide, which disappeared within 30 minutes after irradiation. The dose level had no marked effect on the intensity, the shape, and the rate of disappearance of the observed signals.

The microwave absorption curves of Figs. 3A, 3E, 3I, 3J, and 3K were obtained from pentobarbital, pentobarbital sodium, 1-3-dimethyl pentobarbital, 4,6-dixanthyl pentobarbital, and 4-6-di-*p*-nitrobenzyl pentobarbital respectively, when prepared and irradiated in a boric acid glass. This rigid glassy dispersive medium has been shown to stabilize at room temperature the free radicals produced by radiation in many barbituric acids (Cloutier 1961). The above signals remained unchanged over 10 different experiments and for more than 2 years in 3 of these experiments. It is seen that the same signal, an odd-line spectrum, 110-oersteds total width, was obtained from the irradiated free acid, Fig. 3A; its sodium salt, Fig. 3E; and its dimethyl derivative, Fig. 3I. It is interesting to note that the first two compounds were

TABLE I

Analysis of the gas produced during the irradiation of pentobarbital

Compound	Infrared data		Gas chromatographic data			
			Column A: diisodecyl phthalate		Column J: silica gel	
	Absorption band (microns)	Remarks	Elution time (min)	Relative concentration by volume	Elution time (min)	Relative concentration by volume
Nitrogen	—		0.7	A single peak	0.6	A single peak
Oxygen	—		0.7		0.6	
Carbon monoxide	4.6	Minor component	0.7		0.7	
Methane	7.7	Minor component	0.7		1.0	1
Carbon dioxide	4.3	Major component	0.9	A single peak	9.5	5
Ethane	12.0	Difficult to estimate	0.9		5.7	5
Propane			1.3	1	31	—
Iso-butane			2.3	$\ll 1$		
n-Butane			2.9	<1		
Iso-pentane			5.9	<1		
n-Pentane	3.4	Major component	7.7	8		
n-Hexane			22.0	0		
Water	6.4	Negligible if present				
Hydrogen					—	†
Miscellaneous	—	*				

*No absorption bands corresponding to the following were found at the indicated wavelength in microns: ammonia, 10.4; acetylene, 13.7; allene, 5.1; butadiene, 6.2; cyclopropane, 11.5; ethanol or methanol, 9.5; compounds with carboxyl groups, 6.8; ethylene, 10.5; ethylene oxide, 10.5; hydrogen cyanide, 3.0; ethyl or methyl amine, 12.8; or propylene, 11.0.

†Hydrogen was sought using nitrogen instead of helium as the carrier. No peak corresponding to that of hydrogen was observed. However, column J was not too suitable for the detection of such a light gas.

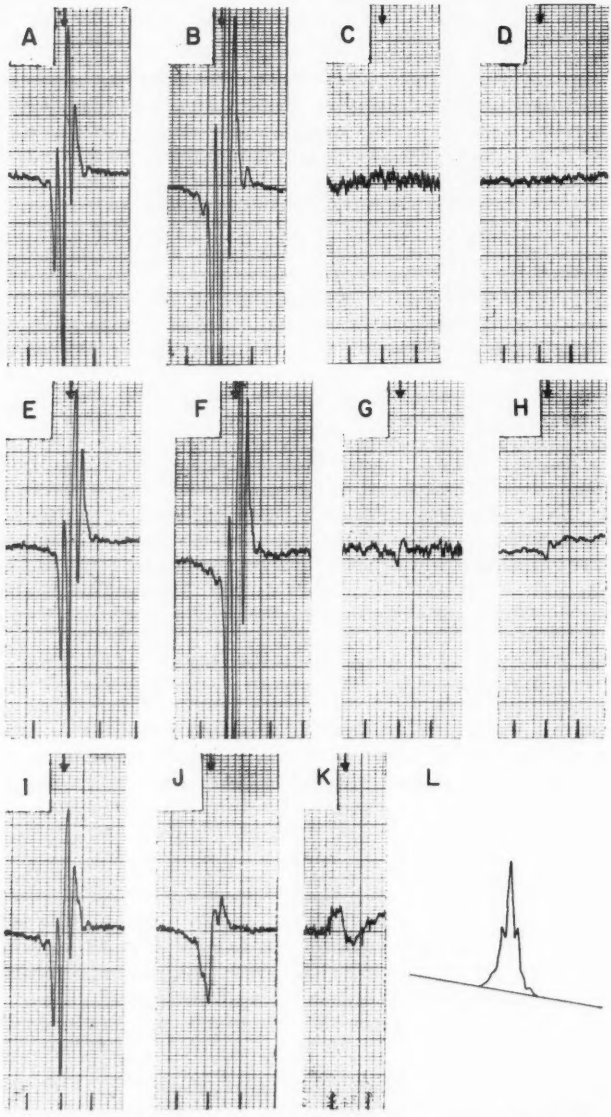


FIG. 3. Paramagnetic resonance absorption signals from gamma irradiated pentobarbital and some of its derivatives at room temperature. First derivative curves. Markers are 50 oersteds apart. Vertical arrows indicate the position D.P.P.H. resonance. Photographs A and B, pentobarbital irradiated in a boric acid glass; photographs C and D, pentobarbital irradiated in air and nitrogen respectively; photographs E and F, pentobarbital sodium irradiated in a boric acid glass; photographs G and H, pentobarbital sodium irradiated in air and nitrogen respectively; photograph I, 1,3-dimethyl pentobarbital irradiated in a boric acid glass; photograph J, 4,6-dixanthyl pentobarbital irradiated in a boric acid glass; photograph K, 4,6-di-*p*-nitrobenzyl pentobarbital irradiated in a boric acid glass; photograph L, the signal of photograph E integrated.

crystalline materials while the last one was an oil at room temperature (Manson and Cloutier 1961). The similarity of the results with solutes of two different physical states is a further experimental proof of the versatility and efficiency of the boric acid glass method of preparation for the stabilization at room temperature of the free radicals produced by radiation in barbituric acids. The xanthyl and *p*-nitrobenzyl derivatives, Figs. 3J and 3K, however, gave an unresolved triplet of about 70 oersteds wide. Qualitatively, the relative importance of the xanthyl or *p*-nitrobenzyl groups on a molecular weight basis changes the configuration of the molecule and would be expected to modify the radiation damage mechanism. These two derivatives were included mainly to supplement the electron paramagnetic resonance study of irradiated pentobarbital but the results were too complex for easy interpretation.

The patterns of Figs. 3A and 3E were repeated at a much higher amplification and are shown in Figs. 3B and 3F respectively. The main structure of the signals was off scale; however, the existence and positions of the weaker outer components were thus ascertained. The signal of Fig. 3E was integrated and is shown in Fig. 3L. The sloping base line was due to the slight asymmetry in the original first derivative curve.

BIOLOGICAL TEST DATA

Fifty Wistar rats, weighing 75 to 114 grams, were sorted randomly according to weight into five equal groups. Pentobarbital sodium, irradiated at 0, 39, 98, 150, and 300 Mrad, was injected subcutaneously, one irradiation level per group, at a dose of 22.5 mg/kg body weight. The drug was administered as a 0.75% solution in water. The dose of 22.5 mg/kg body weight was selected after an elaborate preliminary study of the dose-response characteristics of irradiated and non-irradiated pentobarbital sodium with rats. The criteria used for determining the onset and the end of sleep were based on the reported methods of observation of Moir (1937), Carmichael (1938), and Brook and Cartland (1944).

The average sleeping times observed within groups are listed in Table II

TABLE II

Average sleeping time of rats given a dose of 22.5 mg/kg body weight of irradiated pentobarbital sodium

Radiation dose (Mrad)	Average sleeping time (min)	$\log_e$ of average sleeping time
0	42.6±1.4	3.75185
39	33.4±1.8	3.50856
98	26.4±3.4	3.27336
150	17.2±5.1	2.84491
300	5.3±1.9	1.66771

along with their calculated standard errors. The $\log_e$ of the average sleeping time versus the radiation dose of pentobarbital sodium is plotted in Fig. 4. It was found that, for a dose of 22.5 mg/kg body weight, with rats, a linear

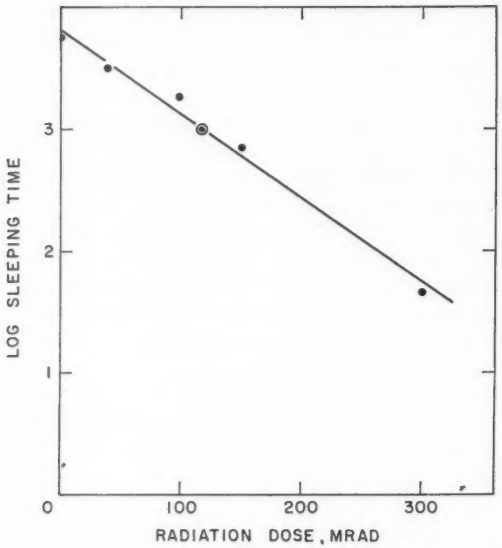


FIG. 4. The logarithm of the average sleeping time of rats given a dose of 22.5 mg/kg body weight of irradiated pentobarbital sodium.

regression of the $\log_e$ of the average sleeping time (T) on radiation dose of pentobarbital sodium (D) was significant at less than 1%, the regression equation being: $\log_e T(\text{min}) = 3.00928 - 0.00697\ D(\text{Mrad}) - 117.4$.

OTHER TESTS DATA

The X-ray powder diffraction patterns of pentobarbital sodium irradiated at 13 progressive levels from 0 to 300 Mrad were obtained. The spectra showed no significant changes in the positions and relative intensities of the lines.

Pentobarbital sodium, irradiated at 0 and 150 Mrad, was assayed by non-aqueous titration to observe any change in potency due to irradiation. The samples were dissolved in 50 ml glacial acetic acid, two drops of crystal violet in CH_3COOH were added, and the solutions were titrated with perchloric acid in dioxane, 0.1 N. Values of 97.8% and 97.2% were obtained from unirradiated and radiated pentobarbital sodium respectively.

An infrared differential absorption spectrum of pentobarbital sodium irradiated at 150 Mrad versus the unirradiated salt was obtained. Minor differences only were indicated on the differential spectrum. These differences could not be interpreted in terms of molecular changes in irradiated pentobarbital sodium.

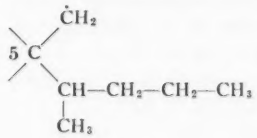
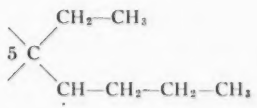
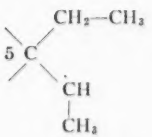
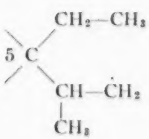
A water solution of 2.68 g of pentobarbital sodium irradiated at 150 Mrad was prepared. An insoluble fraction was isolated by filtration. It represented 3.62% by weight of the original solute. This water insoluble fraction had a broad melting point range whose lower limit was 35.0° C. Its sodium content, as determined by flame photometry, was not significantly different from that of the water soluble fraction.

DISCUSSION

The electron paramagnetic resonance spectra recorded from irradiated pentobarbital, pentobarbital sodium, and 1,3-dimethyl pentobarbital, Figs. 3A, 3E, and 3I respectively, are odd-line spectra. Their contour characteristics indicate hydrogen hyperfine interaction with the electron spin. Further, the observed line intensity ratios in the above spectra strongly suggest that the even number of interacting hydrogen atoms are of the equivalent or the near-equivalent type. The spectra of Figs. 3B and 3F, obtained at a much higher amplification, show that a signal multiplicity as high as nine must be considered even though, in doing so, the line intensity distribution of the signals departs seriously from a true binomial one. However, it is believed that the observed absorption signals are complex. They may result from the superposition of signals of varying odd multiplicities, centered about the same g-value and having the same line spacing (12 oersteds). The relative amounts of each contributing free radical species would be such as to produce the observed line intensity distribution of the signals. Restricting the discussion to the case where a single hit per molecule occurs, the various possibilities with pentobarbital, taking into account the above experimental data and their theoretical implications, are listed in Table III. The solid phase free radicals,

TABLE III

List of the free radicals produced by gamma irradiation of pentobarbital

Solid phase free radicals	Gaseous phase free radicals	Type
	$\dot{C}H_3$	I
	$\dot{C}H_3$	II
	$\dot{C}H_2CH_2CH_3$	III
	$\dot{C}H_2CH_3$	IV
*	$\dot{H}$	V

*H radicals may originate from the destruction of the malonylurea ring structure.

trapped in the boric acid glass matrix, are all expected to give resonance absorption signals which have a line spacing of 12 oersteds (Cloutier 1961). Their *g*-values are also very close to that of the free electron. Therefore, the additive property of any given mixture of such contributing signals to yield the observed absorption patterns must be considered. The gaseous phase free radicals can diffuse and recombine in a number of ways, as illustrated in Table IV. It is interesting to note that Table IV explains the formation of

TABLE IV

Recombination of the gaseous phase free radicals produced by gamma irradiation of pentobarbital

Recombination reaction (types are identified in Table III)	Gas produced
I + V or II + V	CH_4
I + II or IV + V	C_2H_6
III + V	C_3H_8
I + III or II + III or IV + IV	C_4H_{10}
III + IV	C_5H_{12}
III + III	C_6H_{14}

all the gaseous components identified by gas chromatography and infrared spectrophotometry, Table I, with the exception of CO_2 and CO. CO_2 and CO are believed to originate from cases where the malonylurea ring structure of the barbiturate is opened by the incident radiation. The fragments would degenerate rapidly to release some CO or CO_2; any solid residue would be accounted for by the 3.62% water insoluble fraction, as already reported for irradiated pentobarbital sodium. *n*-Hexane, listed in Table IV, was not detected in this experiment. However, it is quite possible that, if produced, its concentration was not high enough to allow positive detection.

The partial destruction by radiation of the side chains in position 5 of pentobarbital sodium, as outlined in Table III, would account for the decrease in average sleeping time values as a function of dose level observed in the biological test with rats. Qualitatively, a reduction in the complexity of the substituents in position 5 decreases the time during which the barbiturate is metabolized by the body, i.e. the length of action of the drug. It is interesting to note that a logarithmic relation between average sleeping time and radiation dose to pentobarbital sodium was found, as would be expected with the present radiation damage model: reduction of average sleeping time directly related to radiation damage which is logarithmically related to radiation dose.

The partial destruction of the side chains in position 5 would also explain the results from the non-aqueous titration of irradiated pentobarbital sodium. The molecules with modified substituents in position 5 would be titrated as barbiturate just as the unmodified one: therefore, little change in potency would be detected. The fragments corresponding to the cases where the malonylurea ring is opened (the 3.62% water insoluble fraction) could either have been titratable as barbiturate or possibly have escaped the sensitivity of the method.

CONCLUSION

If the primary effect of radiation on a molecule is interpreted in terms of a chemical bond rupture produced anywhere in this molecule, irradiation of pentobarbital sodium by gamma or X radiation can either disrupt the ring structure or the side structure in position 5 of the ring. When comparing the results of the biological test and the water solubility test of pentobarbital sodium irradiated at 150 Mrad (60% reduction in average sleeping time with rats as opposed to only 3.62% of water insoluble residue), it would seem evident that radiation damage to the side chain in position 5 is for one reason or another a preferred process. The relative concentration of CO_2, as reported in Table I, appears to contradict this statement. However, of the many recombination processes for the gaseous phase free radicals listed in Table IV, a very important possibility has not been mentioned previously, that is, the capture of slowly diffusing gaseous phase free radicals by neighboring solid phase free radicals. Such an exchange reaction would result in a change in physiological activity of the irradiated drug but not in an increase of the amount of alkane gases produced. Therefore, the relative concentrations of CO_2 and all of the alkane gases, as reported in Table I, may not be truly indicative of the relative frequency of radiation damage to the two molecular areas (ring or side structure).

The site of bond rupture along the side structure of the molecule of pentobarbital is restricted to positions where an even number of nearly equivalent hydrogen atoms is left; therefore, the complete removal by radiation of either one substituent is not to be expected. An immediate consequence is that, since the physiological activity of this type of compound is known to be associated with the 5,5-disubstitution of alkyl, aryl, or alicyclic groups (Levi 1957), radiation, neglecting the less frequent cases where the malonylurea ring is opened, is not expected to destroy the physiological activity of pentobarbital but simply to modify it. There is no definite indication of preferred positions along the two side chains where bond ruptures may be effected. The production of *n*-pentane and ethane has most likely been enhanced by a play of diffusion rates and average recombination times of the gaseous phase free radicals.

ACKNOWLEDGMENTS

The authors are indebted to Mr. C. A. Mainville for the assay of irradiated pentobarbital sodium by non-aqueous titration, to Mr. John H. Hsu for the X-ray diffraction work, and to Mr. J. C. Meranger for technical assistance in the many phases of this project. The 500-curie Transfer Case Irradiator was kindly loaned to this laboratory by the Commercial Products Division of the Atomic Energy of Canada Limited.

REFERENCES

Brook, M. J. and Cartland, C. F. 1944. J. Pharmacol. Exptl. Therap. **80**, 119.
Carmichael, E. B. 1938. J. Pharmacol. Exptl. Therap. **62**, 284.
Chattin, L. G. and Levi, L. 1957. Appl. Spectroscopy, **11**, 177.
Cloutier, J. A. R. 1961. Can. J. Phys. **39**, 514.

CLOUTIER, J. A. R. and MANSON, J. M. 1959. Appl. Spectroscopy, **13**, 34.
FLANN, B. C. and CLOUTIER, J. A. R. 1961. Appl. Spectroscopy. To be submitted.
LEVI, L. 1957. Bull. Narcotics, **9**, 30.
LEVI, L. and HUBLEY, C. E. 1956. Anal. Chem. **28**, 1591.
MANSON, J. M. and CLOUTIER, J. A. R. 1961. Appl. Spectroscopy, **15**, 77.
MCCUTCHEON, R. S. and PLEIN, E. M. 1949. J. Am. Pharm. Assoc. **38**, 24.
MOIR, W. 1937. J. Pharmacol. Exptl. Therap. **59**, 68.
PIERSON, R. H., FLETCHER, A. N., and GANZT, E. S. 1956. Anal. Chem. **28**, 1218.

CHANGES IN THE DIURNAL HOUR OF MAXIMUM OF THE COSMIC-RAY INTENSITY[1]

J. KATZMAN

ABSTRACT

The diurnal hour of maximum of the meson component changed progressively at Ottawa, Canada, from 10 hr 44 min to 14 hr 40 min during the period January 1955 to December 1960 while the nucleon component changed from 12 hr 12 min to 15 hr 16 min for the same period. This evidence favors the 22-year cycle in the diurnal hour of maximum that was first suggested by Thambyahpillai and Elliot, for stations within a geomagnetic latitude belt between 58.1° N. and 48.1° S. The diurnal hour of maximum at Churchill changed from 14 hr 40 min to 15 hr 24 min during the period April 1957 to December 1960 for the meson component and from 15 hr 12 min to 15 hr 52 min for the nucleon component. Although the change was for a later hour the indication of a 22-year cycle at Churchill is not impressive. At Resolute the diurnal hour of maximum is dominated by the varying magnetic masses in interplanetary space. It is shown that the anisotropy varies both in magnitude and direction depending on the conditions that exist in the solar system.

INTRODUCTION

The intensity of the primary cosmic-ray particles arriving at the earth is not completely isotropic. A small anisotropy of about 0.3% in solar time does exist at stations not within the Arctic circle and in order to determine the source or cause of this anisotropy, the solar daily variation has been studied extensively by many authors.* The possible existence of a 22-year cycle in the diurnal hour of maximum for the period 1932–1952 similar to the 22-year cycle in the polarity of sunspots was demonstrated by Thambyahpillai and Elliot (1953) for eight stations in a geomagnetic latitude belt between 58.1° N. and 48.1° S. An attempt was made by Brunberg and Dattner (1954) to determine the general direction of the anisotropy for these stations after correcting for the deflection of the primary particles in the earth's dipole field. They concluded that the daily variation is caused by an excess of particles arriving from the 18-hour direction and is due to the tangential component of the solar wind. No attempt was made to explain the hour of maximum obtained at Manchester, which does not agree with this conclusion. If the cycle in the diurnal hour of maximum is repeated in the present sunspot polarity cycle at a station other than Manchester then the above explanation by Brunberg and Dattner of the diurnal hour of maximum is not adequate. Ottawa and Manchester are at the same geomagnetic latitude and hence comparable.

The present paper deals with the results obtained at Ottawa, Churchill, and Resolute. Ottawa (56.8° N.) is within the latitude belt mentioned above,

[1]Manuscript received June 5, 1961.
Contribution from the Division of Pure Physics, National Research Council, Ottawa, Canada.
Issued as N.R.C. No. 6498.

*See "Cosmic Ray Variations" by L. I. Dorman, State Publishing House for Technical and Theoretical Literature, Moscow 1957.

Churchill (68.7° N.) is further north, and Resolute (82.9° N.) is a polar station. At Resolute the cutoff magnetic rigidity is effectively zero and the magnetic lines of force move out into space beyond 10 earth radii and become entangled with whatever magnetic fields exist in interplanetary space. Thus if an anisotropy always exists in the low energy portion of the primary particle spectrum it will be apparent at Resolute at a time of low magnetic turbulence in interplanetary space.

Clark *et al.* (1961) have shown that primary particles of energies $>10^{15}$ ev are isotropic. An anisotropy that is due to the sun will be due to particles of much lower energies and a change in the diurnal hour of maximum can be expected with changes in the sun's activity. Also, the solar dipolar field which was considered fixed in polarity is now known to reverse its polarity (Babcock 1961). It is important, therefore, to determine whether the 22-year cycle in the diurnal hour of maximum is genuine and, if so, this will add weight to the hypothesis that the daily variation is extraterrestrial.

It will be shown that there is strong evidence at Ottawa that the beginning of the 22-year cycle in the diurnal hour of maximum is repeating itself in the present sunspot polarity cycle but that the effect at Churchill is not impressive. There is no indication of a 22-year cycle at Resolute. There is evidence, however, that the diurnal hour of maximum at Resolute changes with changes in the planetary magnetic index K_p.

EXPERIMENTAL RESULTS

The cosmic-ray intensity as measured by neutron monitors is influenced by the changes in the atmospheric pressure, whereas the meson component is influenced by the distribution of the mass of air above the recording instrument. This mass distribution will depend upon the time of year and hence if 12-month moving averages are used this effect will be negligible.

Twelve-month moving averages of the pressure corrected data from cubical telescopes and neutron monitors at Ottawa, Churchill, and Resolute were analyzed harmonically. The amplitude and phase of the first harmonics of the meson component are shown in Fig. 1. It can clearly be seen that the diurnal hour of maximum at Ottawa becomes progressively later as the solar cycle progresses but does not follow the activity of the sun. The hour of maximum at Churchill changes by 44 minutes while for the same period the change is 2 hours 16 minutes at Ottawa. The diurnal hour of maximum at Resolute appears to move through a cycle.

The amplitude and phase of the nucleon component for the three stations are shown in Fig. 2. In 1955, the neutron monitor at Ottawa contained two counters. The design of the pile was changed and the new pile contained three larger counters (Fenton, Fenton, and Rose 1958). The diurnal hour of maximum for 1955 is given but not included in the 12-month moving averages which were started with the data from the new pile. It can be accepted that the diurnal hour of maximum of the nucleon component followed that of the meson component between 1955 and 1956 since the hour of maximum changed by 1 hour 52 minutes for the nucleon component and by 1 hour 36 minutes

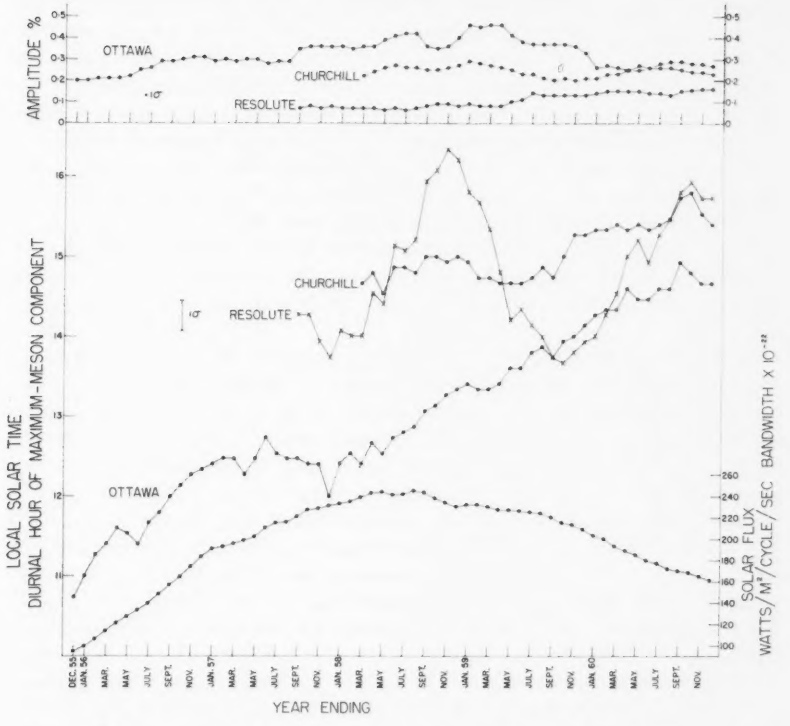


FIG. 1. Twelve-month moving averages of the amplitude and hour of maximum of the first harmonic of the meson component and the activity of the sun as measured by the solar 10.7-cm radio flux, supplied by Covington.

for the meson component and in both cases the change was towards a later hour. The same apparent cycle is present in the nucleon component at Resolute.

DISCUSSION

The daily variation of the cosmic-ray intensity has been considered to be due to an anisotropy of the cosmic rays and the rotation of the earth. Since the earth possesses a magnetic field and if this field only is considered then the diurnal hour of maximum will vary as one moves from the equator to the pole, due to the deflection of the charged primaries by the earth's field. Hence, if the anisotropy, the magnetic field, and rotation of the earth are the only factors influencing the diurnal hour of maximum then the hour of maximum should remain constant at any particular station provided there is no change in the magnetic rigidity spectrum of the particles and no significant change in the earth's magnetic field. That the cosmic-ray intensity changes with solar activity has been amply demonstrated by Forbush (1954) for two solar

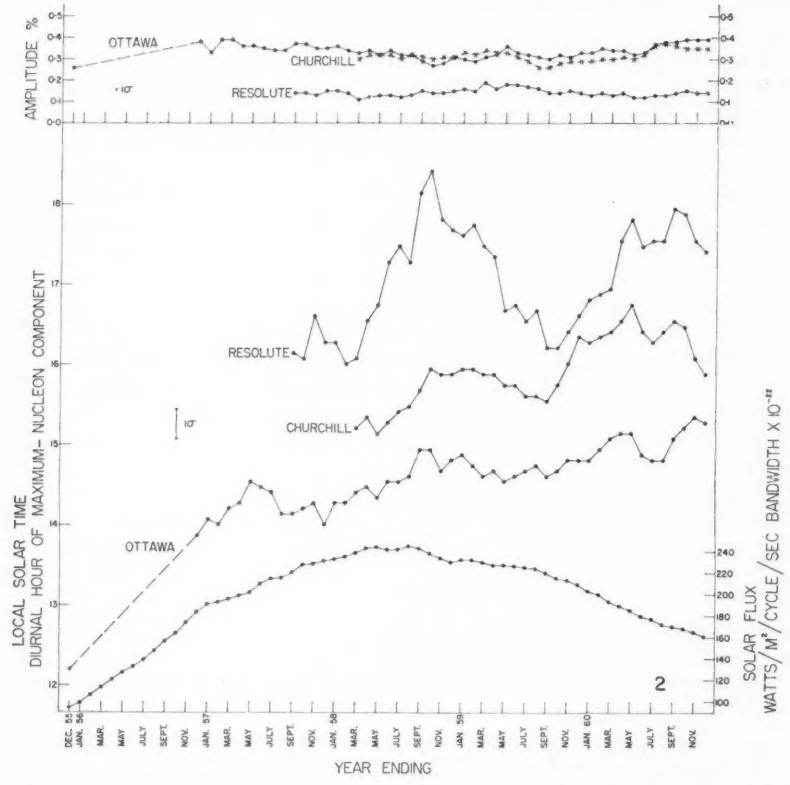


FIG. 2. Twelve-month moving averages of the amplitude and hour of maximum of the first harmonic of the nucleon component and the activity of the sun as measured by the solar 10.7-cm radio flux, supplied by Covington.

cycles. The earth's field, though disturbed, can be considered constant for the short period under study. The intensity of the low magnetic rigidity particles is reduced with increase in solar activity and, therefore, the diurnal hour of maximum should either remain constant or change to earlier hours.

Thambyahpillai and Elliot (1953) have shown that the change in the diurnal hour of maximum for the meson component has an approximate 22-year cycle similar to the 22-year cycle in the polarity of sunspots. For approximately eight years, 1939 to 1946, the hour of maximum was at 15 ± 0.5 hr and at either end of the 22-year cycle at about 11 hr 30 min. The hour of maximum for the meson component at Ottawa changed from 10 hr 44 min to 14 hr 40 min from January 1955 to December 1960. The change in the diurnal hour of maximum was progressive with no significant abrupt change. The nucleon component changed from 12 hr 12 min to 15 hr 16 min. The difference in the hour of maximum between the two components is

negligible for 1960 but very significant for 1955. Both components reached the same diurnal hour of maximum at approximately 15 hours at a time of decreasing solar activity though the intensity of the nucleon component has not increased significantly above the low at the time of maximum solar activity. Since it is considered that the intensity of low magnetic rigidity particles is decreased thus causing the decrease in the general intensity, the diurnal hour of maximum should not progress to later hours without a corresponding increase in intensity. Hence it is obvious that the anisotropy is not directly influenced by the earth's magnetic field at Ottawa. The direction of the primary particles must be influenced by fields of force external to the earth's magnetic field and any anisotropy would be due to the resultant of all the fields. This resultant field will be different for different parts of the world.

The diurnal hour of maximum at Resolute of the nucleon and meson components differ continuously by approximately two hours. The apparent cycle which, for this epoch, appears to have a period of almost two years is present in both components and in the planetary magnetic index K_p with the minima and maxima preceding by three months those of the cosmic-ray diurnal hour of maximum, Fig. 3. Since the cycle in the hour of maximum does not appear at either Ottawa or Churchill it must be concluded that the diurnal hour of maximum at Resolute is more sensitive to the interplanetary magnetic disturbances as measured by the planetary magnetic index K_p.

To differentiate between the mechanisms that influence the daily variation it is important to determine (1) whether or not the diurnal hour of maximum remains in the midafternoon at stations within a geomagnetic latitude belt between approximately 60° N. and 60° S., at the next minimum of solar activity, (2) whether or not the anisotropy at stations near the geomagnetic poles is controlled continuously by the magnetic state of interplanetary space. If these conditions prevail then the anisotropy is caused by a prime mechanism with two different effects (i) an anisotropy related to the 22-year cycle in the general polarity of the sun and observed at stations within a geomagnetic latitude belt between 60° N. and 60° S., (ii) an anisotropy at polar stations controlled by some magnetohydrodynamic effect in interplanetary space.

To determine the diurnal hour of maximum at Resolute at a time of low interplanetary disturbances, the data from a second telescope stationed there were analyzed. This telescope is similar to the one used aboard the Canadian icebreaker "Labrador" (Rose and Katzman 1955) in measuring cosmic rays during a cruise through the Northwest Passage. The amplitude and phase of the first harmonic is shown in Fig. 4 together with the data from the I.G.Y. cubical telescope. The gap in the data from the smaller "Labrador telescope" is due to faulty operation of the equipment and to the fact that the 12-month moving averages were started from March 1956. This telescope was shut down in early January 1960 when it was found that trouble had developed sometime earlier. The diurnal hour of maximum for the year November 1954 to October 1955 that was observed, at Resolute, when K_p was very low, was at 02 hr with an amplitude of only $0.026 \pm 0.013\%$. We place very little

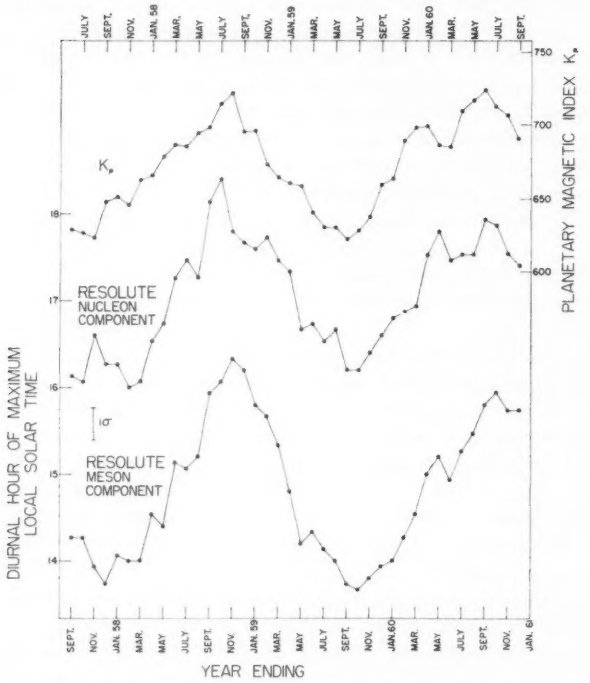


FIG. 3. Twelve-month moving averages of the diurnal hour of maximum of the meson and nucleon components at Resolute and the monthly mean of the 12-month moving averages of the planetary magnetic index K_p shifted by 3 months.

reliance on phase angles when the amplitude is so low. The amplitudes obtained from the data of both telescopes are the same for the period when both were operating; it is permissible, therefore, to accept the result of the smaller telescope for the earlier period. Also the hours of maximum follow each other fairly well except for the last months when sporadic trouble developed in the smaller telescope. Hence it is considered that, at a time of low interplanetary magnetic disturbance, if an anisotropy does exist at Resolute in the meson component it would be of the order of $0.02 \pm .01\%$ at 02 hr. With increasing interplanetary magnetic disturbance as measured by K_p the intensity of the anisotropy increases at Resolute and the hour of maximum changes.

The amplitude of the diurnal variation of the meson component at Ottawa increased by 100% from 1955 to 1958 and then decreased to approximately the 1955 value (Table I) while the 1960 value of the nucleon component is 50% above the 1955 value. At Churchill the changes in the amplitudes are not very significant. At Resolute, though the amplitude of the first harmonic of the nucleon component remains constant, the amplitude of the meson component increased by 100% between 1957 and 1960. If the magnetic rigidity spectrum is simply depressed during the time of solar maximum

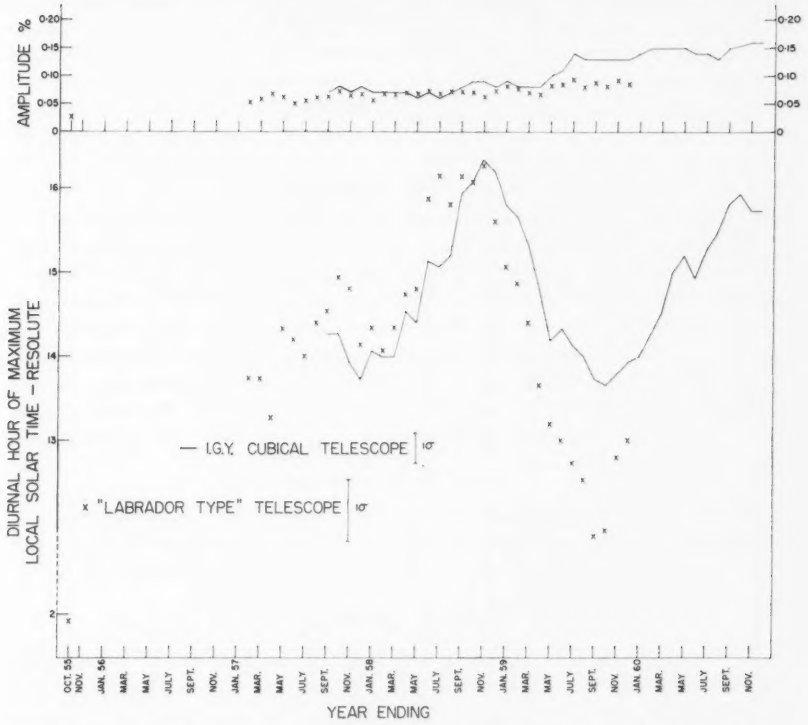


FIG. 4. Twelve-month moving averages of the amplitude and diurnal hour of maximum of the meson component as obtained from the data of the I.G.Y. and "Labrador Type" telescopes at Resolute.

activity, as discussed by Simpson (1960), then the amplitude of the diurnal variation should remain constant. This is approximately true for the nucleon component but not for the meson component. Since particles of all energies can arrive at Resolute and the air mass is the only large absorber, then the amplitude of the nucleon component should be higher than the meson component, due to the decay of the meson. This is so for 1957 and 1958 but not for 1959 and 1960 when the amplitudes are equal.

If it is accepted that at Resolute the diurnal hour of maximum is modulated by the interplanetary magnetic masses, it is permissible to obtain a mean value for the period October 1956 to December 1960. The mean diurnal hour of maximum for the meson component is at 15 hr ±20 min and at 17 hr 8 min ±24 min for the nucleon component. At Churchill, the mean hour of maximum for the meson component is at 15 hr 4 min and at 15 hr 48 min for the nucleon component. At Ottawa, the hour of maximum for both components becomes progressively later and for 1960 the values are 14 hr 40 min and 15 hr 12 min for the meson and nucleon components respectively. It is obvious from the

TABLE I

Annual values of the amplitude and phase of the first harmonics of the meson and nucleon components of the cosmic-ray intensity at Ottawa, Churchill, and Resolute

		1955		1956		1957		1958		1959		1960	
Station	Component	Amplitude, %	Phase, degrees	Amplitude, %	Phase, degrees	Amplitude, %	Phase, degrees	Amplitude, %	Phase, degrees	Amplitude, %	Phase, degrees	Amplitude, %	Phase, degrees
Ottawa	Meson	0.20	161	0.31	$\pi+5$	0.36	π	0.40	$\pi+20$	0.33	$\pi+32$	0.23	$\pi+40$
	Nucleon	0.26	$\pi+3$	0.38	$\pi+28$	0.35	$\pi+30$	0.31	$\pi+42$	0.33	$\pi+42$	0.40	$\pi+48$
Churchill	Meson					0.22*	$\pi+36$	0.27	$\pi+45$	0.21	$\pi+49$	0.27	$\pi+51$
	Nucleon					0.33*	$\pi+47$	0.32	$\pi+58$	0.29	$\pi+65$	0.35	$\pi+58$
Resolute	Meson					0.08	$\pi+26$	0.08	$\pi+63$	0.13	$\pi+29$	0.16	$\pi+56$
	Nucleon					0.15	$\pi+64$	0.15	$\pi+85$	0.14	$\pi+69$	0.14	$\pi+81$

*Data from April to December 1957.

above discussion that the direction of the anisotropy cannot be obtained by considering only the deflection of the primary particles caused by the earth's field. The anisotropy varies both in magnitude and direction depending on the conditions in the solar system. As more information about these conditions becomes known, a clearer picture will emerge with regard to the cause of the observed anisotropy.

CONCLUSIONS

The diurnal variation of the cosmic-ray intensity at Ottawa, Churchill, and Resolute is not consistent with an anisotropy that is fixed in space or time. The hour of maximum of the diurnal variation at Ottawa becomes progressively later, which is an indication that the diurnal variation follows the 22-year cycle in the polarity of sunspots as was first demonstrated by Thambyahpillai and Elliot. The anisotropy of the meson component at Resolute is questionable at a time of low planetary magnetic index K_p but is definite for high values of K_p when the direction of the anisotropy becomes dominated by the interplanetary magnetic turbulence.

ACKNOWLEDGMENTS

It is a pleasure to acknowledge the contributions of the staff of the Cosmic Ray Group of the Division of Pure Physics of the National Research Council which has enabled this study.

REFERENCES

BABCOCK, H. W. 1961. Astrophys. J. **133**, 572.
BRUNBERG, E. A. and DATTNER, A. 1954. Tellus, **6**, 73.
CLARK, G. W., EARL, J., KRAUSHAAR, W. L., LINSLEY, J., ROSSI, B. B., SHERB, F., and SCOTT, W. D. 1961. Phys. Rev. **122**, 637.
FENTON, A. G., FENTON, K. B., and ROSE, D. C. 1958. Can. J. Phys. **36**, 824.
FORBUSH, S. E. 1954. J. Geophys. Research, **59**, 525.
ROSE, D. C. and KATZMAN, J. 1956. Can. J. Phys. **34**, 1.
SIMPSON, J. A. 1960. Astrophys. J. Suppl. No. 44, **4**, 378.
THAMBYAHPILLAI, T. and ELLIOT, H. 1953. Nature, **171**, 918.

HIGH-FREQUENCY DIFFRACTION BY A SPHERE[1]

C. L. TANG[2]

ABSTRACT

A systematic procedure is given for the determination of the asymptotic series directly from the Helmholtz equation and the boundary conditions for the field in the shadow region of a sphere illuminated by a plane wave at high frequencies. The first two terms in the series for the shadow region, including the regions near the axial caustic and the boundary layer near the surface of the sphere, are explicitly evaluated. The present procedure can be generalized to any smooth convex three-dimensional object with a rotational symmetry, illuminated by a plane wave in the direction of the axis of rotational symmetry.

1. INTRODUCTION

In recent times the problem of high-frequency diffraction by obstacles of "complex shape" whose geometry does not fit into a simple co-ordinate system has received considerable attention. Keller (Levy and Keller 1957) has treated the problem of diffraction by a convex smooth object using a geometrical theory of diffraction that is a generalization of the geometrical optics of the reflected and refracted rays through the introduction of additional rays called the "diffracted" or the "surface" rays. The method of Keller provides one more term in various asymptotic developments in addition to that given by geometrical optics. But, without further elaboration, no information about the higher-order terms can be obtained by the above method. In a recent paper (Wu 1957), a new method has been proposed by Wu by which it is possible to obtain higher-order terms in the asymptotic series for the field in the shadow region of a circular cylinder through the use of the boundary layer theory. This procedure was later used by Wu and Seshadri (1958) to solve the problem of diffraction of plane waves by an infinite cylinder whose cross section is a smooth convex curve. Together with the method of Keller *et al.* (1956), which works satisfactorily for the illuminated region, the problem of diffraction by an infinite convex cylinder can be considered to be well understood. The next problem in line for investigation is the diffraction by a convex three-dimensional obstacle with a rotational symmetry, illuminated by a plane wave in the direction of the axis of rotational symmetry. Again, in the illuminated region the method of Keller can be used satisfactorily. In the shadow region, there appears an additional feature which is not present in the case of a cylinder; that is, an axial caustic appears in the shadow region along the axis of rotational symmetry of the body. In this report the simplest case of this class of obstacles, the sphere, is solved in the sense that a systematic procedure is given for the determination of the asymptotic series for the field in the shadow region, including the regions near the axial caustic and the surface of the sphere. The generalization to the case of an arbitrary convex

[1]Manuscript received May 26, 1961.
Contribution from the Cruft Laboratory, Harvard University, Cambridge, Massachusetts.
[2]Now at Research Division, Raytheon Company, Waltham, Massachusetts.

body with a rotational symmetry will be given in a subsequent report. Although the explicit evaluation for the latter case is considerably more complicated, there are no new basic difficulties in principle.

2. STATEMENT OF THE PROBLEM AND THE INTERIOR SOLUTION OF THE SHADOW REGION

The problem to be studied is as follows. For the incident field

$$\psi^{\text{inc}} = \exp(-ikz), \tag{2.1}$$

it is desired to determine the asymptotic behavior, as k tends to infinity, of the total field ψ which satisfies the time-reduced Helmholtz equation

$$L\psi = (\Delta + k^2)a^2\psi = 0 \tag{2.2}$$

and the Dirichlet boundary condition

$$\psi = 0 \tag{2.3}$$

on the surface of the sphere $r = a$.

Define the physical space external to the sphere as $r \geqslant a$, $0 \leqslant \phi \leqslant \pi$, and $0 \leqslant \theta \leqslant 2\pi$, where $\theta = 0$ (or 2π) is the positive z-axis. Because of symmetry, it is sufficient to consider the behavior of ψ in the positive shadow region, in the sense of geometrical optics, for a given value of ϕ.

The variables to be used in the study of the field in the shadow region, except for the region near the axial caustic, are u and t, defined for a given ϕ, as follows (Fig. 1):

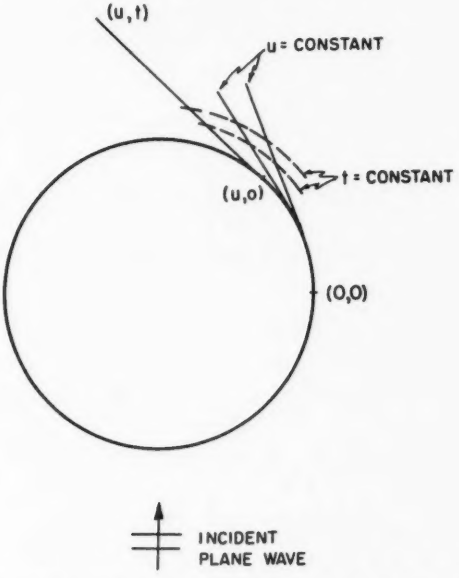


FIG. 1. The (u,t) co-ordinate system for the shadow region.

$$t = (r^2-a^2)^{\frac{1}{2}}/a,$$
$$u = \theta-\frac{\pi}{2}-\cos^{-1}\frac{a}{r}. \tag{2.4}$$

The origin of the co-ordinate system is chosen to be the point of intersection of the positive shadow boundary with the boundary curve of the sphere for the given ϕ.

In the present co-ordinates, the wave equation takes the following form:

$$L\psi = (\Delta+k^2)a^2\psi = \left\{\left(1+\frac{1}{t^3}\right)\frac{\partial^2}{\partial t^2}+\frac{2t^2-1}{t^3}\frac{\partial}{\partial t}-\frac{2}{t^2}\frac{\partial^2}{\partial t\partial u}+\left[\frac{1}{t^3}-\frac{1}{t}+\frac{1}{t-\cot u}\right]\frac{\partial}{\partial u}+\frac{1}{t^2}\frac{\partial^2}{\partial u^2}+(ka)^2\right\}\psi = 0. \tag{2.5}$$

The general solution ψ is to be split into a slowly varying amplitude factor and a rapidly oscillating exponential factor T, where the form of T,

$$T = \exp[ika(u+t) + i(ka)^{\frac{1}{3}}f(u)], \tag{2.6}$$

is determined from the partial wave expansion.

From the commutation relations

$$\left[\frac{\partial}{\partial t}, T\right] = ikaT \tag{2.7a}$$

and

$$\left[\frac{\partial}{\partial u}, T\right] = [ika+i(ka)^{\frac{1}{3}}f']T, \tag{2.7b}$$

it follows that

$$T^{-1}LT = ika\left[2\frac{\partial}{\partial t}+\frac{1}{t}+\frac{1}{t-\cot u}\right]-(ka)^{2/3}\frac{f'^2}{t^3}+i(ka)^{\frac{1}{3}}\left\{f'\left[\frac{2}{t^2}\left(\frac{\partial}{\partial u}-\frac{\partial}{\partial t}\right)+\frac{1}{t^3}-\frac{1}{t}+\frac{1}{t-\cot u}\right]+f''\right\}+\Delta a^2. \tag{2.8}$$

From (2.6) and (2.8), the following form is supposed for the wave function in the shadow region:

$$\psi = T\sum_{n=0}^{\infty}F_{sn}(u,t)(ka)^{-n/3}. \tag{2.9}$$

When (2.8) is applied to (2.9), the following results are obtained:

$$F_{s0} = C_{s0}(u)\frac{1}{\sqrt{t(t-\cot u)}}, \tag{2.10}$$

$$F_{s1} = \left[\frac{if'^2}{2t}C_{s0}(u)+C_1(u)\right]\frac{1}{\sqrt{t(t-\cot u)}}, \tag{2.11}$$

where the constants of integration $C_{s0}(u)$ and $C_{s1}(u)$ depend on u.

Equations (2.10) and (2.11) and the F_{sn}'s determined by continuing the iterative procedure are not valid in the vicinities of $t = 0$ or/and $t = \cot u$, corresponding to the regions near the surface of the sphere and the axial caustic. Consider first the region near the axial caustic.

3. THE FIELD NEAR THE AXIAL CAUSTIC

In view of the fact that $\partial/\partial u$ in the (u,t) co-ordinate system, in the neighborhood of the axial caustic, is taken along the direction in which the field varies most rapidly, the following change of co-ordinate is suggested (Fig. 2).

$$\xi = t - \cot u, \tag{3.1}$$

$$u = u.$$

In the (u,ξ) co-ordinates, in the neighborhood of the axial caustic, all the rapid variations in ψ and its partial derivatives are now in ξ only.

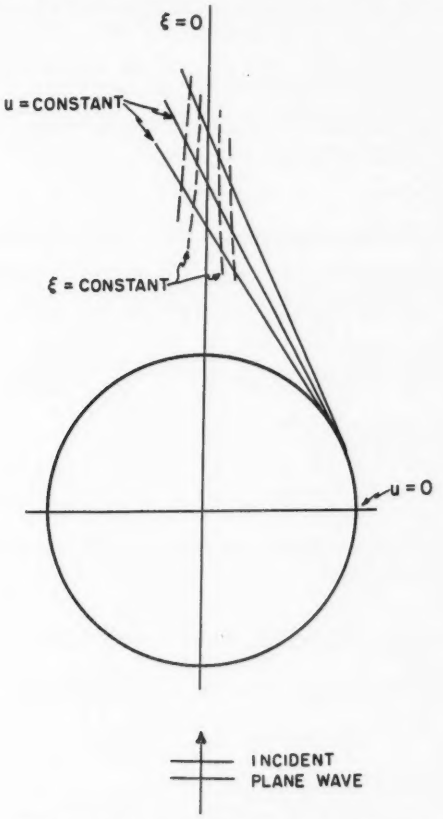


FIG. 2. The (u,ξ) co-ordinate system for the axial caustic region.

With (3.1), (2.8) becomes

$$
\begin{aligned}
(3.2)\qquad T^{-1}LT = {} & ika\left[2\frac{\partial}{\partial\xi}+\frac{1}{\xi+\cot u}+\frac{1}{\xi}\right]-(ka)^{2/3}\frac{f'^2}{(\xi+\cot u)^2} \\
& +i(ka)^{\frac{1}{3}}\left\{\frac{2f'\cot^2 u}{(\xi+\cot u)^2}\frac{\partial}{\partial\xi}+\frac{2f'}{(\xi+\cot u)^2}\frac{\partial}{\partial u}\right. \\
& \left.+f'\left[\frac{1}{(\xi+\cot u)^3}-\frac{1}{\xi+\cot u}+\frac{1}{\xi}\right]+\frac{if''}{t^2}\right\} \\
& +\left\{\left[1+\frac{\cot^4 u}{(\xi+\cot u)^2}\right]\frac{\partial^2}{\partial\xi^2}+\left[\frac{2(\xi+\cot u)^2-1}{(\xi+\cot u)^3}\right.\right. \\
& \left.+\left(\frac{1}{(\xi+\cot u)^3}-\frac{1}{\xi+\cot u}+\frac{1}{\xi}\right)(1+\cot^2 u)\right. \\
& \left.+\frac{2(1+\cot^2 u)}{(\xi+\cot u)^2}\frac{\partial}{\partial u}-\frac{2\cot u(1+\cot^2 u)}{(\xi+\cot u)^2}\right]\frac{\partial}{\partial\xi} \\
& \left.+\left[\frac{1}{(\xi+\cot u)^3}-\frac{1}{\xi+\cot u}+\frac{1}{\xi}\right]\frac{\partial}{\partial u}+\frac{1}{(\xi+\cot u)^2}\frac{\partial^2}{\partial u^2}\right\}.
\end{aligned}
$$

In accordance with the spirit of the boundary layer theory (Carrier 1953) an examination of (3.2) indicates that the change of variable

$$(3.3)\qquad \eta = (ka)\xi$$

gives the proper co-ordinate stretching. With (3.3), (3.2) becomes

$$(3.4)\qquad T^{-1}LT = (ka)^2 L_{1c} + (ka)^{5/3}L_{2c} + (ka)^{4/3}L_{3c} + 0[(ka)]$$

where

$$(3.5)\qquad L_{1c} = (1+\cot^2 u)\frac{\partial^2}{\partial\eta^2}+\left[2i+\frac{1+\cot^2 u}{\eta}\right]\frac{\partial}{\partial\eta}+\frac{i}{\eta},$$

$$(3.6)\qquad L_{2c} = 0,$$

$$(3.7)\qquad L_{3c} = 2if'\frac{\partial}{\partial\eta}+\frac{if'}{\eta}.$$

In view of (2.9), the following form is supposed for the wave function in the region near the axial caustic

$$(3.8)\qquad \psi_c = T\sum_{n=0}^{\infty} F_{cn}(u,\eta)(ka)^{-n/3}.$$

Therefore the first three partial differential equations in the caustic region are

$$(3.9)\qquad L_{1c}F_{c0} = 0,$$

$$(3.10)\qquad L_{1c}F_{c1} + L_{2c}F_{c0} = 0,$$

$$(3.11)\qquad L_{1c}F_{2c} + L_{2c}F_{c1} + L_{c3}F_{c0} = 0,$$

and the boundary conditions are that all the F_{cn}'s must be bounded for $\eta = 0$ and that ψ_c must blend with the interior solution, ψ, in the regions where both solutions predict the same behavior of ψ. From (3.9) and the boundary conditions, it follows that

$$F_{c0} = C_{s0}(u)\frac{\sqrt{2ka\pi}\sin u\, e^{-i\eta\sin^2 u}}{2\sqrt{\cot u}\cos \pi/4} J_0(\eta \sin^2 u). \tag{3.12}$$

From (3.10), (3.12), and the appropriate boundary conditions, it follows that

$$F_{c1} = \left[\frac{if'^2}{2\cot u} C_{s0}(u)+C_{s1}(u)\right]\frac{\sqrt{2ka\pi}\sin u\, e^{-i\eta\sin^2 u}}{2\sqrt{\cot u}\cos \pi/4} J_0(\eta \sin^2 u). \tag{3.13}$$

This iterative procedure can be continued to give all the higher-order terms. The $C_{sn}(u)$'s remain to be determined.

4. THE FIELD IN THE BOUNDARY LAYER NEAR THE SURFACE OF THE SPHERE; CONCLUDING REMARKS

The field in the region near the boundary curve of the sphere and the various constants of integration, $C_{sn}(u)$, in (2.10) and (2.11) can again be determined through the use of the boundary layer theory and the Dirichlet boundary condition. Let

$$t_1 = (ka)^{\frac{1}{3}}t. \tag{4.1}$$

Then a rearrangement of terms on the right-hand side of (2.8) gives

$$\begin{aligned} T^{-1}LT = (ka)^{4/3}&\left\{\frac{1}{t_1^3}\frac{\partial^2}{\partial t_1^2}+\left(2i-\frac{2if'}{t_1^2}-\frac{1}{t_1^3}\right)\frac{\partial}{\partial t_1}+\left(\frac{i}{t_1}+\frac{if'}{t_1^3}-\frac{f'^2}{t_1^2}\right)\right\} \\ &+(ka)\left\{\left(-\frac{2}{t_1^2}\frac{\partial}{\partial u}\right)\frac{\partial}{\partial t_1}+\frac{if''}{t^2}-\frac{i}{\cot u}+\left(\frac{2if'}{t_1^2}+\frac{1}{t_1^3}\right)\frac{\partial}{\partial u}\right\} \\ &+(ka)^{2/3}\left\{\frac{\partial^2}{\partial t_1^2}+\frac{2}{t_1}\frac{\partial}{\partial t_1}-\frac{it_1}{\cot^2 u}-\frac{if'}{t_1}+\frac{1}{t_1^2}\frac{\partial^2}{\partial u^2}\right\}+0[(ka)^{\frac{1}{3}}]. \end{aligned} \tag{4.2}$$

The leading term in (4.2) can be simplified by letting

$$T_1 = \exp[-i(\tfrac{1}{3}t_1^3-f't_1)]. \tag{4.3}$$

Then it follows that

$$T_1^{-1}T^{-1}LTT_1 = (ka)^{4/3}L_a+(ka)L_b+(ka)^{2/3}L_c+0[(ka)^{1/3}] \tag{4.4}$$

where

$$L_a = \left(\frac{1}{t_1}\frac{\partial}{\partial t_1}\right)^2-2f'+t_1^2, \tag{4.5}$$

$$L_b = \left(-\frac{2}{t_1^2}\frac{\partial}{\partial t_1}+2i+\frac{1}{t_1^3}\right)\frac{\partial}{\partial u}-\frac{i}{\cot u}+\frac{if''}{t^2}, \tag{4.6}$$

$$L_c = \left[\frac{\partial^2}{\partial t_1^2} - \frac{it_1}{\cot^2 u} - \frac{if'}{t_1} + \frac{1}{t_1^2}\frac{\partial^2}{\partial u^2}\right] + \left[-(t_1^2 - f')^2 + 2\left(\frac{1}{t_1} - it_1^2 + if'\right)\frac{\partial}{\partial t_1} - \frac{2}{t_1}(2it_1^2 - if')\right]. \tag{4.7}$$

Let

$$\psi = TT_1F_T(u,t_1;k) \tag{4.8}$$

where

$$F_T(u,t_1;k) = \sum_{n=0}^{\infty} F_{Tn}(u,t_1)(ka)^{-n/3}. \tag{4.9}$$

Once F_{Tn} is known, C_{sn} and hence F_{sn} and F_{cn} can be determined by the asymptotic expansion of F_{Tn}. This will be illustrated for the case $n = 0$.

The application of (4.4) to (4.8) and (4.9) gives the first three partial differential equations in the boundary layer:

$$L_aF_{T0} = 0, \tag{4.10}$$

$$L_aF_{T1} + L_bF_{T0} = 0, \tag{4.11}$$

$$L_aF_{T2} + L_bF_{T1} + L_cF_{T0} = 0. \tag{4.12}$$

The outgoing-wave solution of (4.10) is the following Airy function.

$$F_{T0}(t,u) = C_{T0}(u)\mathrm{Ai}[e^{-i(\pi/3)}2^{-2/3}(t_1^2 - 2f')]. \tag{4.13}$$

The application of the Dirichlet boundary condition to (4.13) gives

$$f' = M = (2.3381074)2^{-\frac{1}{3}}\exp(i\pi/3). \tag{4.14}$$

From (4.11) and the boundary condition

$$F_{T1}(u,0) = 0$$

it follows that

$$\int_0^{\infty} \mathrm{Ai}[e^{-i(\pi/3)}2^{-2/3}(t_1^2 - 2M)]\left[C'_{T0}(u)\left(2i + \frac{1}{t_1^3} - \frac{2}{t_1^2}\frac{\partial}{\partial t_1}\right) - C_{T0}(u)\frac{i}{\cot u}\right] \times \mathrm{Ai}[e^{-i(\pi/3)}2^{-2/3}(t_1^2 - 2M)]t_1 dt_1 = 0.$$

Through integration by parts, and since the integral of the square of the Airy integral is not zero, it follows that

$$\frac{\partial}{\partial u}C_{T0}(u) = \frac{C_{T0}(u)}{2\cot u}, \tag{4.15}$$

or

$$C_{T0}(u) = \frac{C_0}{\sqrt{\cos u}}. \tag{4.16}$$

Therefore

$$F_{T0}(u,t) = \frac{C_0}{\sqrt{\cos u}} \mathrm{Ai}[e^{-i(\pi/3)} 2^{-2/3}(t_1^2 - 2M)]. \tag{4.17}$$

Thus the leading term in the asymptotic development for the field in the boundary layer near the sphere is completely determined.

Before proceeding to the determination of $F_{T1}(t,u)$, it is now interesting to determine $C_{s0}(u)$ and to complete the leading term of the interior solution and the solution near the axial caustic. From (4.17), using the asymptotic representation for the Airy integral and by comparison with (2.10), the following result is obtained.

$$C_{s0}(u) = C_0 \pi^{-1/2} e^{i(7\pi/12)} 2^{-5/6} (ka)^{-1/6} (\sin u)^{-1/2}. \tag{4.18}$$

From (4.18) and (2.10), the following interior solution is obtained.

$$\psi = C_0 \pi^{-1/2} e^{i(7\pi/12)} 2^{-5/6} T(ka)^{-1/6} \{[t(t - \cot u)\sin u]^{-1/2} + 0[(ka)^{-1/3}]\}. \tag{4.19}$$

From (4.18) and (3.12), the following solution near the axial caustic region is obtained.

$$\psi_c = C_0 \pi^{-1/2} e^{i(7\pi/12)} 2^{-5/6} T(ka)^{-1/6} \left\{ \frac{\sqrt{2ka\pi}\sin u}{2\sqrt{\cos u}\cos \pi/4} \right. \\ \left. \exp[-ika \sin^2 u(t - \cot u)] J_0[ka \sin^2 u(t - \cot u)] + 0[(ka)^{-1}] \right\}. \tag{4.20}$$

Two remarks are now in order. First, the functional forms of ψ and ψ_c as given in (4.19) and (4.20) are in agreement with results of either the partial wave expansion or Keller's (1957) geometrical theory of diffraction. Secondly, the fact that the phase decreases by $\pi/2$ as a ray passes through a focus is placed in evidence by the factor $(\cos u)^{-1/2}$ in (4.17), since $\cos u$ changes sign as u passes a focus ($u = (2n + 1/2)\pi$, where $n = 0, 1, 2, \ldots$).

For the second term of the asymptotic series for the field in the shadow region, it is only necessary to determine F_{T1}. From (4.12), (4.17), (4.18), and the boundary condition

$$F_{T2}(u,0) = 0,$$

it follows that

$$\frac{\partial}{\partial u} C_{T1}(u) - \frac{C_{T1}(u)}{2\cot u} + \frac{C_0}{\sqrt{\cos u}} \left[\frac{iM}{2}(M-2) + \frac{iI_3}{2I_1}(1-2M) \right. \\ \left. + \frac{iI_5}{2I_1} - \frac{I_2}{4I_1} \frac{1}{\cot^2 u} \right] = 0 \tag{4.21}$$

where

$$I_n = \int_0^\infty t_1^n \{\mathrm{Ai}[e^{-i(\pi/3)} 2^{-2/3}(t_1^2 - 2M)]\}^2 dt_1. \tag{4.22}$$

The integral I_n is summable in the sense of Abel and hence is a number (Wu and Seshadri 1958). The solution of (4.21) gives

$$C_{T1}(u) = \frac{C_1}{\sqrt{\cos u}} - \frac{C_0}{\sqrt{\cos u}}[(\alpha+\beta)u - \beta \tan u] \tag{4.23}$$

where

$$\alpha = \frac{iM}{2}(M-2) + \frac{iI_3}{2I_1}(1-2M) + \frac{iI_5}{2I_1}, \tag{4.24}$$

$$\beta = \frac{I_2}{2I_1}. \tag{4.25}$$

Therefore

$$F_{T1}(u,t) = \frac{C_0}{\sqrt{\cos u}}\left[\frac{t_1}{\cot u} - u(\alpha+\beta) + \beta \tan u\right] \times \mathrm{Ai}[e^{-i(\pi/3)}2^{-2/3}(t_1^2-2M)] + \frac{C_1}{\sqrt{\cos u}}\mathrm{Ai}[e^{-i(\pi/3)}2^{-2/3}(t_1^2-2M)]. \tag{4.26}$$

The second term in the asymptotic series for the field in the shadow region is thus completely determined.

Several remarks should be made in connection with the results obtained. Two normalization constants C_0 and C_1 are still undetermined, although these are not of prime importance here. This is due to the fact that the magnitude of the incident wave has not been used. They can be determined through comparison with a partial wave expansion at any one point in the shadow region.

Secondly, the results are not valid in the neighborhood of the foci ($t = \cot u = 0$). The situation here is more complicated, since the boundary layer theory does not apply. The size of the focus region is $\sim 0[(ka)^{-1}]$ centered around any focus. Hence it goes to zero as ka tends to infinity.

ACKNOWLEDGMENTS

The author is indebted to Professor R. W. P. King for his encouragement and for criticizing the manuscript. He would also like to acknowledge support by an IBM Fellowship during the period in which this work was completed.

REFERENCES

CARRIER, G. F. 1953. Advances in Appl. Mech. **3**, 1.

KELLER, J. B., LEWIS, R. M., and SECKLER, B. D. 1956. Commun. Pure Appl. Math. **9**, 207.

LEVY, B. R. and KELLER, J. B. 1957. Diffraction by a Smooth Object, Research Report No. EM-109, Institute of Mathematical Sciences, New York University (December).

WU, T. T. 1957. The Electromagnetic Theory of Light. I, Scientific Report No. 9, Cruft Laboratory, Harvard University (April).

WU, T. T. and SESHADRI, S. R. 1958. The Electromagnetic Theory of Light. II, Scientific Report No. 22, Cruft Laboratory, Harvard University (November).

UNCHARGED CONDUCTING TOROIDAL RING IN A UNIFORM ELECTRIC FIELD[1]

S. C. Loh[2]

ABSTRACT

Mathematical expressions for the potential function of an uncharged conducting toroidal ring placed in a uniform electric field are derived and expressed in terms of toroidal functions. Some numerical results were calculated by the IBM 650 computer at the University of Toronto and are included in the present paper. To verify the calculated results, a systematic study of an electrolytic tank was undertaken. It was found that the theoretical calculations agreed well with the experimental results.

INTRODUCTION

The potential functions of an uncharged conducting ring in a uniform electric field are usually calculated approximately, assuming that the diameter of the conductor is small compared with that of the ring. In the present paper the exact solutions of the problem have been derived by means of toroidal functions. Two cases have been discussed, namely, (i) a uniform field of electric force parallel to the axis of the ring and (ii) a uniform electric field perpendicular to the axis of the ring. Therefore, by the principles of superposition the general potential functions of an uncharged conducting ring placed in a uniform electric field in any direction may be derived. Some numerical results calculated by IBM 650 computer at the University of Toronto are presented in order to illustrate the derived solutions.

A. A UNIFORM ELECTRIC FIELD PARALLEL TO THE AXIS OF TORE

Consider a conducting ring of radius R, as shown in Fig. 1. The ring is at zero potential and is situated in a field that had a uniform static field intensity E_0 before the ring was introduced. That is, the original potential distribution is given by

$$\Phi = E_0 z. \tag{1}$$

This must also be the distribution for large z after the ring is introduced. The problem, then, is to find the potential Φ everywhere outside the ring, the potential and field being zero inside the tore.

The potential function Φ of an electrostatic field system always satisfies Laplace's equation. The general solution of Laplace's equation in toroidal co-ordinates is given by Carter and Loh (1958) as follows:

[1]Manuscript received May 24, 1961.
Contribution from the Radio and Electrical Engineering Division, National Research Council, Ottawa, Canada, and the Department of Electrical Engineering, University of Hong Kong, Hong Kong.
Issued as N.R.C. No. 6505.
[2]National Research Council Postdoctorate Fellow 1957–1959. Present address: Electrical Engineering Department, University of Hong Kong, Hong Kong.

$$\Phi = \sqrt{(\cosh u - \cos v)} \sum_{n=0}^{\infty} \{a_n p_n(u) + b_n q_n(u)\} \cos(nv + \alpha_n) \tag{2}$$

where a_n, b_n, and α_n are arbitrary constants,

$p_n(u)$ and $q_n(u)$ are known as "toroidal functions" of the first and second kinds respectively.

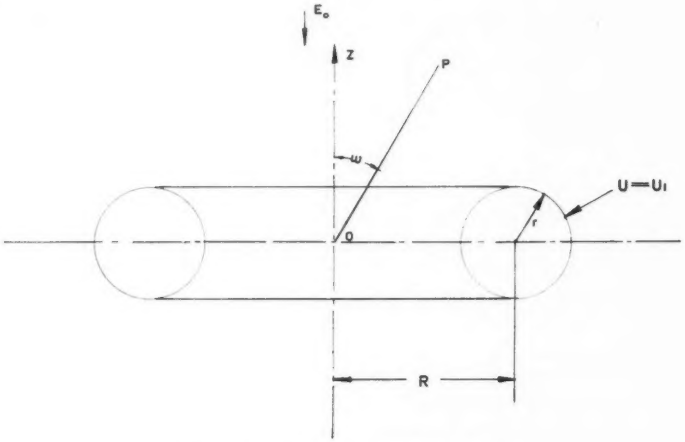


FIG. 1. Cross section of conducting ring.

The toroidal co-ordinates (u, v, w) are related to the well-known cylindrical polar co-ordinates (ρ, ϕ, z) by the following expressions:

$$\begin{cases} \rho = \dfrac{a \sinh u}{\cosh u - \cos v}, \\ \phi = w, \\ z = \dfrac{a \sin v}{\cosh u - \cos v}. \end{cases} \tag{3}$$

The relations between a, R, r, and u_1 are

$$u_1 = \text{arc sech } (r/R),$$

$$a = \sqrt{(R^2 - r^2)}.$$

When the system is symmetrical with the central plane, as in the present case, $\alpha_n = 0$; therefore the general potential function becomes

$$\Phi = \sqrt{(\cosh u - \cos v)} \sum_{n=0}^{\infty} \{a_n p_n(u) + b_n q_n(u)\} \cos nv. \tag{4}$$

The arbitrary constants a_n and b_n must be so chosen as to satisfy the boundary conditions, namely, (a) $\Phi = 0$ where $u = u_1$ for all values of v and (b) $\Phi = E_0 z$ when z approaches infinity. We thus get

$$a_n = \frac{-2\sqrt{2}}{\pi} a E_0 . \delta_n . n \frac{q_n(u_1)}{p_n(u_1)} \frac{\sin nv}{\cos nv} \tag{5}$$

and

$$b_n = \frac{2\sqrt{2}}{\pi} a E_0 . n . \delta_n . \frac{\sin nv}{\cos nv} \tag{6}$$

where $\delta_n = 1$ when $n \geqslant 1$, but $\delta_0 = \frac{1}{2}$.

Therefore the complete solution for the potential function Φ is

$$\Phi = \frac{2\sqrt{2}}{\pi} E_0 a \sqrt{(\cosh u - \cos v)} \sum_{n=0}^{\infty} n . \delta_n \left\{ q_n(u) - \frac{q_n(u_1)}{p_n(u_1)} p_n(u) \right\} \sin nv. \tag{7}$$

The density at any point of the tore is

$$\sigma = \epsilon_0 \left(\frac{1}{h_n} \frac{\partial \Phi}{\partial u} \right)_{u=u_1}, \tag{8}$$

where h_n is the scale factor of the toroidal co-ordinate system,

$$= \epsilon_0 \frac{(\cosh u_1 - \cos v)}{a} E_0 a \frac{2\sqrt{2}}{\pi} \left\{ \sqrt{(\cosh u_1 - \cos v)} \sum_{n=0}^{\infty} n . \delta_n \left[q_n'(u_1) - \frac{q_n(u_1)}{p_n(u_1)} p_n'(u_1) \right] \right\}.$$

But

$$q_n'(u_1) - \frac{q_n(u_1)}{p_n(u_1)} p_n'(u_1) = 1/[\sinh u_1 . p_n(u_1)]. \tag{9}$$

Thus

$$\sigma = \frac{2\sqrt{2}}{\pi} \epsilon_0 E_0 \frac{(\cosh u_1 - \cos v)^{3/2}}{\sinh u_1} \sum_{n=0}^{\infty} \left\{ n \delta_n . \frac{\sin nv}{p_n(u_1)} \right\}. \tag{10}$$

B. A UNIFORM ELECTRIC FIELD PERPENDICULAR TO THE AXIS OF THE TORE

The general potential function derived in (7) is no longer valid for the present case where the electric field is perpendicular to the axis of the ring, since this is a three-dimensional problem. However, the general solution of Laplace's equation for the toroidal co-ordinate system (u, v, w) has been derived by the author (1959) as follows:

$$\Phi = \sqrt{(\cosh u - \cos v)} \sum_n \sum_m \{ (a_m \cos mw + b_m \sin mw) \times (c_n \cos nv + d_n \sin nv) \times [A_{mn} p_n^m(u) + B_{mn} q_n^m(u)] \} \tag{11}$$

where a_m, b_m, c_n, d_n, A_{mn}, and B_{mn} are arbitrary constants. When ρ is large, the field is undisturbed and must be equal to

$$\Phi = E_0 \rho \cos w = E_0 a \frac{\sinh u \cos w}{\cosh u - \cos v}. \tag{12}$$

For the potential Φ to vary as in (12) for large values of ρ, where the field is undisturbed by the presence of the ring, the coefficients b_m, d_n, and B_{mn} must be equal to zero and also $A_{mn} = 0$ for all values of m except $m = 1$. Thus the potential function becomes

$$\Phi = E_0 \rho \cos w + \sqrt{(\cosh u - \cos v)} \cos w \sum_{n=0}^{\infty} A_n p_n^1(u) \cos nv \tag{13}$$

where A_n is a new constant.

Introducing the boundary condition that

$$\Phi = 0 \tag{14}$$

$$\text{at } u = u_1 \text{ for all values of } v, \tag{15}$$

we have from (13) that

$$0 = E_0 a \frac{\sinh u_1 \cos w}{\cosh u_1 - \cos v} + \sqrt{(\cosh u_1 - \cos v)} . \cos w \sum_{n=0}^{\infty} A_n p_n^1(u_1) \cos nv. \tag{16}$$

In order to evaluate the arbitrary constant A_n, we multiply eq. (16) by $\cos nv$ and integrate from 0 to 2π. Equation (16) yields to

$$\frac{\pi}{2} A_n p_n^1(u_1) + \delta_n E_0 a \sinh u_1 \int_0^{\pi} \frac{\cos nv \, dv}{(\cosh u_1 - \cos v)^{3/2}} = 0 \tag{17}$$

where $\delta_n = 1$ when $n \geqslant 1$ but $\delta_0 = \frac{1}{2}$.

The integral may be evaluated readily by using the following identity:

$$\int_0^{\pi} \frac{\cos nv}{(\cosh u - \cos v)^{3/2}} dv = \frac{-2\sqrt{2}}{\sinh u} q_n'(u) \tag{18}$$

where $q_n'(u)$ are the first derivative of $q_n(u)$. Substituting (18) into (17), we get

$$A_n = \frac{4\sqrt{2}\delta_n E_0 a}{\pi p_n^1(u_1)} q_n'(u_1). \tag{19}$$

Hence the general expression for the electric potential at any point (u, v) outside the ring $(u = u_1)$ having a uniform electric field perpendicular to the axis of the ring is given by

$$\Phi = E_0 a \frac{\sinh u \cos w}{\cosh u - \cos v} + \sqrt{(\cosh u - \cos v)} . \cos w \sum_{n=0}^{\infty} \frac{4\sqrt{2}\delta_n E_0 a}{\pi p_n^1(u_1)} q_n'(u_1) p_n^1(u) \cos nv. \tag{20}$$

As before, the densities of charge induced on the ring can be, if necessary, readily evaluated.

ELECTROLYTIC TANK INVESTIGATION

In order to verify the calculated results an electrolytic tank investigation was carried out. As a toroidal ring is an axially symmetric body a wedge-

shaped tank was used with the apex of the wedge representing the axis of the toroidal ring. A brass disk electrode was placed such that its distance from the wedge apex was equal to the toroidal radius R. The probe used was of 0.010 in. diameter platinum wire. An oscilloscope was used for null indication.

A preliminary test was carried out to determine the accuracy obtainable with the experimental setup. The experimental equipotential lines for an electrode system representing concentric cylinders were compared with the known exact solution. It was concluded that the experimental error was within 1%.

CONCLUSIONS

The potential function Φ derived in eq. (7) for the electric field parallel to the axis of the ring has been calculated, using the first four terms of the series, for two cases with different ratio of R/r, by the electronic computer IBM 650 at the University of Toronto. Figures 2 and 3 show the equipotential

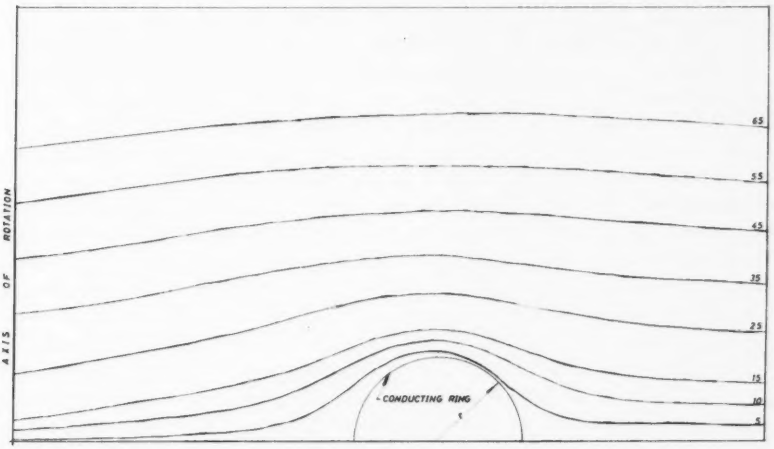


FIG. 2. Uncharged conducting ring in originally uniform field ($R/r = 5$).

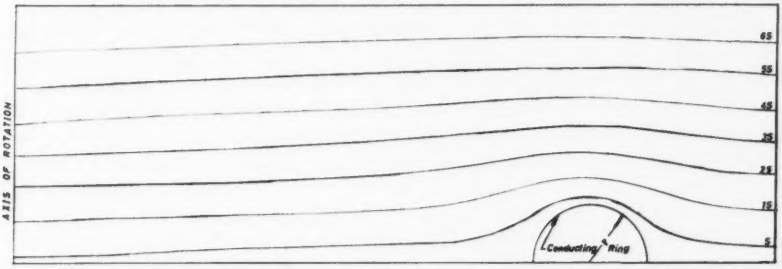


FIG. 3. Uncharged conducting ring in originally uniform field ($R/r = 10$).

lines near the ring which is placed in a uniform electric field parallel to its axis. It is found that the potential field at a distance $z > 2r$ is nearly undisturbed by the presence of the ring and is independent of the ratio R/r. However, the field near the ring is dependent on the ratio of R/r. The calculated values agreed closely with experiment (1%).

Though the solutions for the potential function have been obtained separately for two cases, namely (i) the field parallel to the axis of the ring and (ii) the field perpendicular to the axis of the ring, the general potential function for the ring placed in a uniform electric field in any direction may be derived by means of the principle of superposition.

ACKNOWLEDGMENT

The author wishes to thank the National Research Council of Canada for the award of a Postdoctorate Fellowship. The problem was initiated whilst the author was holding the fellowship.

REFERENCES

CARTER, G. W. and LOH, S. C. 1958. Proc. Inst. Elect. Engrs. C, **105**, 13.
LOH, S. C. 1959. Can. J. Phys. **37**, 619, 698.

SUBSTRUCTURES AND DISLOCATIONS PRODUCED DURING SOLIDIFICATION OF ALUMINUM[1]

HERALDO BILONI

ABSTRACT

The substructure and the dislocation arrangement produced during polycrystalline solidification of 99.99% aluminum has been studied using the decanting method for the interruption of the solidification. Etching techniques to reveal individual dislocations and the method of thin epitaxial layers of aluminum oxide to reveal changes in impurity concentration have been employed simultaneously. The development of the structure towards the interior of the solid was studied by successive electrolytic polishings. In this way it was possible to show that the "cell substructure" formed by dislocations originates in areas where "microsegregation" of impurities has occurred. Towards the interior of the solid a dislocation rearrangement into a "macromosaic" substructure which is closely related to the cellular substructure occurs. At the same time a substructure of "macrosegregation" occurs, arising from the original grain boundaries. The results are interpreted in the light of recent theories of solidification.

I. INTRODUCTION

Chalmers and his collaborators (Rutter and Chalmers 1953; Tiller and Rutter 1956) have described the formation of substructures during solidification of metals which are closely linked to the impurity content (corrugated and dendritic substructures) and have explained their origin as a result of "constitutional supercooling" (Winegard and Chalmers 1954).

On the other hand, Teghtsoonian and Chalmers (1951) have pointed out that a macromosaic substructure is formed in monocrystals during solidification which is composed of crystalline zones which are slightly disoriented with respect to each other and which have great stability even after prolonged annealing. These two characteristics indicate that the boundaries of macromosaics are formed as a result of the alignment of dislocations, as happens in the case of boundaries of subgrains in polygonized substructures.

The theories which have been put forward most frequently to explain the formation of dislocations during the process of solidification are:

(*a*) Nucleation and ulterior collapse of vacancies (Frank 1955). Schoeck and Tiller (1960) have proved, however, that this mechanism does not suffice by itself to explain the formation of macromosaic or striated substructures.

(*b*) Microsegregation of solute. Tiller (1958), developing ideas originally expressed by Friedel (1956), supposes that in the stages of the advancing solid–liquid interface (Burton *et al.* 1957) and in the borders of the cell substructure, the microsegregations of solute produce variations in the crystalline lattice parameter. These distortions are eliminated through an adequate formation and distribution of dislocations. During the process of cooling they rearrange themselves to minimize the energy, thus forming the macromosaic substructure.

[1]Manuscript received April 19, 1961.
Contribution from the Metallurgical Division, National Atomic Energy Commission, Buenos Aires, Argentine Republic.

It is the purpose of the present work to study the relation existing between cell substructure and the formation of dislocations, and their subsequent rearrangement during the process of cooling. The method used is a combination of the technique of etch pits and of a method using thin epitaxial layers (Biloni and Destaillats 1959), which were deposited by anodic oxidation.

II. EXPERIMENTAL METHODS

The samples were melted in an alumina crucible and cast into a nuclear graphite mold of square section. Prior to completion of solidification, the remaining liquid was drained, exposing the solid–liquid interface (Walton and Chalmers 1959). Viewed under the microscope and without previous preparation the interface presents an appearance similar to that in the works of Chalmers and his collaborators, that is, of cells whose borders are at a lower level than their centers as shown in Fig. 1.

The test tubes were electrolytically polished with Jacquet reagent (1956), composed of 333 cc perchloric acid and 666 cc acetic anhidride, with a voltage of 25 v and electrolytic refrigerated vat.

For the production of etch pits, the reagent recommended by Wyon, Marchin, and Lacombe (1959) was adopted; the composition used was the following: HNO_3, 49% (d = 1.39); HCl, 10% (d = 1.185); butylcellosolve, 40%; AuCl, 10 p.p.m.; HF, 1%.

The method of preparation of the reagent and the procedures followed were those recommended by the above-mentioned authors. With this method both micro and geometrical etch pits are developed. This fact, however, does not affect the results. The reagent is used only once and the period of attack is of about one hour. Nevertheless, careful control of the time is necessary in each case, since the presence of impurities in the bath has substantial effects upon the time of attack (Wyon *et al.* 1959).

In order to obtain conclusions from the data furnished by the techniques of corrosion figures, it must be kept in mind that in the case of aluminum (Lacombe and Wyon 1954) (*a*) etch pits reveal the atmospheres formed by certain solutes around dislocations; (*b*) for 99.99% aluminum, the works of Wyon and Lacombe (1954) prove in a semiquantitative manner that the responsible atom is the Fe atom.

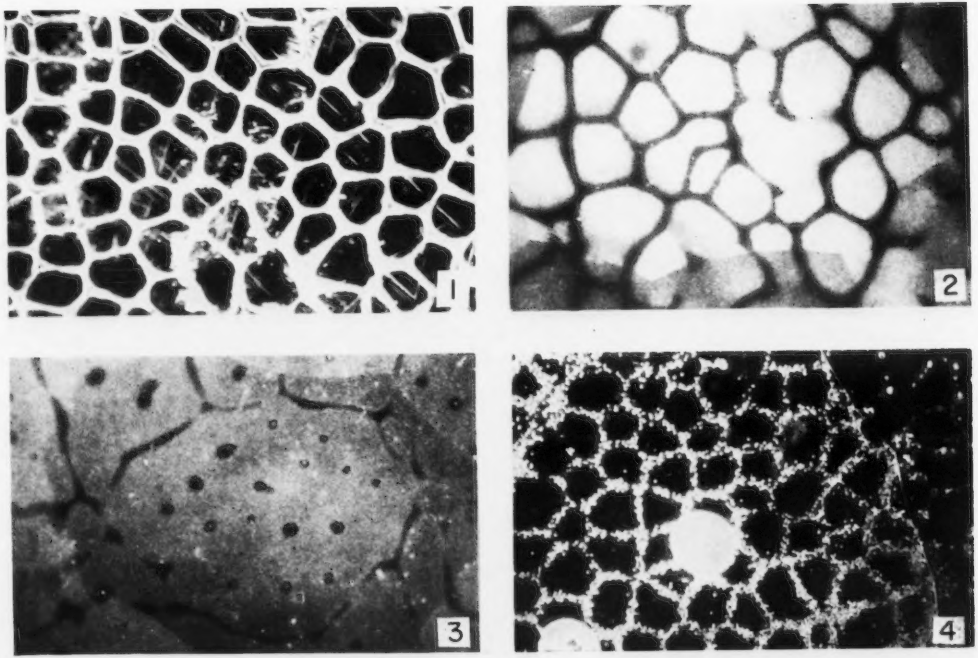
The detection of segregation substructures was carried out by applying the method of thin epitaxial layers that show interference colors under oblique light (Lacombe and Mouflard 1953). In the case of aluminum, anodic oxidation

FIG. 1. Solid–liquid interface resulting from the interruption of solidification, when draining the remaining liquid. Cellular substructure with borders enriched in solute. ×80.

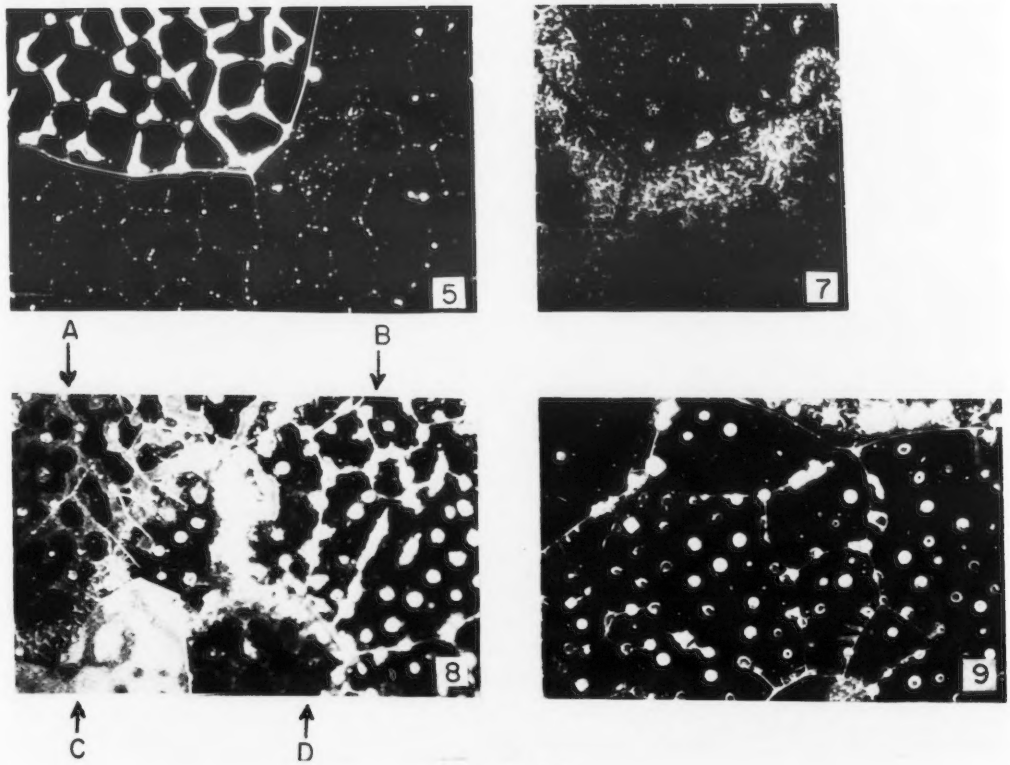
FIG. 2. Solid–liquid interface slightly polished with Jacquet reagent and anodically oxidated in order to form thin epitaxial layers. Cellular substructure may be detected by the different color of the borders compared with the centers, a result of the concentration of solute. It may be seen that one cell shares two or more grains, because of the migration of the grain boundary during the process of cooling. ×80.

FIG. 3. Same zone as shown in Fig. 2, situated 50 microns towards the inner core of the grain. It may be seen that the epitaxial layers only indicate the persistence of the vertices of the cells. The continuous line corresponds to the so-called macrosegregation, a survival of the original size of the grain. ×80.

FIG. 4. Interface slightly polished with Jacquet reagent and attacked with fluorated aqua regia. Formation of etch pits along borders of cellular substructure. ×80.



FIGS. 1–4.



FIGS. 5 7, 8 9.

was used in a 5% solution of H_2SO_4, with a tension of 30 v. The period of time varies between 30 and 60 seconds, but since it is additive, it is possible to control it perfectly well in order to obtain maximum contrast in coloration.

In order to study the relation between etch pits and segregation substructures, the sample attacked with fluorated aqua regia reactive was subjected to anodic oxidation under conditions of operation identical with those described.

All micrographs were taken in dark field illumination, thus contributing to the better detection of corrosion microfigures and to a remarkable improvement of the formation of interference colors.

III. RESULTS OF THE EXPERIMENT

A. Segregation Substructures

We have limited our study to substructures of cellular character, also called Smialowski substructures, the interfaces of which present the aspect shown in Fig. 1.

Electrolytic polishing of a few microns, followed by anodic oxidation, makes it possible to determine, by the difference in the interference colors, the concentration of solute along the border of the cells (Biloni and Destaillats 1959). The method of epitaxial layers indicates differences of orientation in the crystalline structure, since the thickness of the layers is determined by this orientation. This may be appreciated in Fig. 2, as also the superposition of the cellular network of the Smialowski substructure, which may be detected by the different coloration compared with the matrix.

The explanation for the detection of segregated areas by means of the method of thin epitaxial layers is not yet established, and the possible mechanisms are two:

(*a*) Variation of the lattice parameters of the matrix due to impurities can change the kinetics of anodic oxidation and lead to the formation of different thicknesses of the layers (Schoeck, private communication).

FIG. 5. Zone of the interface prepared like the one shown in the preceding figure. Formation of etch pits along the borders of the cells. It may be observed that the response to the chemical attack varies in each grain. In the one shown in the upper part, the attack revealed segregated areas besides producing micro etch pits. ×80.

FIG. 7. Zone situated 200 microns from the interface, attacked with Wyon and Lacombe reagent for several minutes. Concentration of etch pits in the vertices of the cells and zones corresponding to macrosegregation. ×80.

FIG. 8. Zone situated several hundred microns away from the interface, electrolytically polished, attacked with fluorated aqua regia reactive, and then anodically oxidated. Grain A presents a strong, heavy concentration of etch pits in the vertices of the cells and alignments of microfigures that delimit the macromosaic substructure, closely linked with the borders of the cells detected by epitaxial layers. Grain B presents concentration of micro etch pits in the vertices of the cells and a macromosaic substructure delimited by alignment of micro etch pits, which present different sizes though always passing through the vertices of the cells. The concentration of micro pits appearing in the center of the picture corresponds to a macrosegregated region comprising grains A, B, and D. Grains C and D have suffered practically no attack from the fluorated aqua regia reactive and the epitaxial layers reveal different stages of detection of the cell substructure. ×80.

FIG. 9. Zone close to the former, but situated 100 microns towards the interior of the grains. The sample has been electrolytically polished and attacked with fluorated aqua regia reagent. Concentration of etch pits in the vertices of the cells, which has produced overattack. Micro etch pits aligned macromosaic substructure, generally passing through cell vertices. ×80.

(*b*) Different local composition can cause the difference in thickness of the layers, which in this case would result from a chemical reactivity different from that of the matrix.

It is known that each grain grows with a type of cell substructure the morphology of which depends upon crystalline orientation (Tiller and Rutter 1956). In Fig. 2, it is possible to observe the presence of two or more grains that share the same substructure. Of course, in this case, when the border of a cell passes across a grain boundary, the color on the two sides of the latter is different: to the action of segregation is added that of different crystalline orientation. These facts are explained if we take into account the possibility of migration of the grain boundary during the period of cooling after solidification.

If the electrolytic polishing is continued towards the interior of the grains, the appearance of the segregation in the borders of the cells changes after several 10-micron polishings. While the cellular substructure persists in some grains, in others the borders of the cells disappear and only the vertices remain. There is a wide scale of variation between these two cases, and there is always an intensification of color in the cell vertices. This is interpreted as an indication of the fact that when there has been sufficient homogenization, the detection of segregations through thin epitaxial layers depends to a large extent upon the orientation of the surface of each grain. The persistence of the vertices of the cells is explained by there being regions where three hexagons meet, i.e., segregation is greater than along the sides.

Macrosegregation

Figure 3 illustrates the existence of a substructure consisting of continuous segregated zones, connected with the grain boundaries of the substructure and forming a net within which there is a whole number of cells. On the other hand, analysis of Fig. 2 makes it possible to verify that the actual boundaries of the grain have attained their position through migration during the process of cooling.

We conclude that the borders of the macrosegregation substructure correspond to the limits of the cells which originally formed the boundary of the grain.

During the process of cooling, the mobility of the grain boundary has been large enough to produce migration to the present positions, but segregation has remained as a proof of their original position.

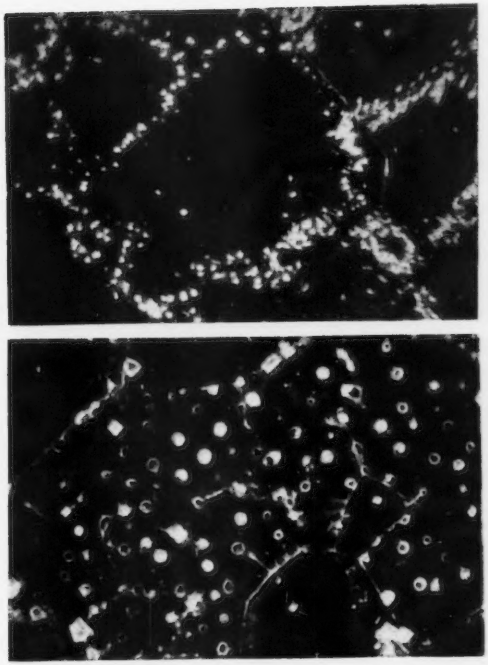
B. Relation between Etch Pits and Segregation Substructures

The purpose of the superposition of the techniques of etch pits and of thin epitaxial layers is to study the relation between segregation substructures and dislocations existing in the metal at the time of solidification, taking

FIG. 6. Same zone as shown in Fig. 3. Brief attack with fluorated aqua regia followed by anodic oxidation. Geometrical etch pits concentrated along borders of macrosegregation substructure and in some cell vertices.

FIG. 10. The same zone as that of the preceding picture, but subjected to anodic oxidation following attack of etch pits. Owing to the interference colors it may be seen that while the vertices of the cells are perfectly clear (because of either their color or the concentration of etch pits which have produced overattack), the borders are much more diffuse. Besides, it may be observed that the alignments of etch pits generally pass through the vertices of the cells but not necessarily through their borders.

PLATE III



into special account that the etch pits detect dislocation-associated impurities. Our experimental results correspond to the successive application of both techniques or to the utilization of etch pits only, and this point is clearly indicated in each case. In this work we have studied the distribution of etch pits starting from the solid–liquid interface and their relation with segregated areas.

(a) Solid–Liquid Interface

The interface, slightly polished with Jacquet reagent and attacked with fluorated aqua regia reagent, presents a strong concentration of etch pits in the borders of the cells of the Smialowski substructure (Figs. 4 and 5). It may be observed in the figures that the response to the chemical attack is the result of the crystalline orientation of the grains, and after a prolonged attack even the substructure of segregation was revealed in one of them (Fig. 5).

(b) Variation of Relation between Segregated Zones, and Etch Pits According to Distance from the Interface

Just as the aspect of segregated zones changes according to the distance from the interface, the same happens with the distribution of etch pits. At several tens of microns we observe a strong concentration in the macrosegregated zones and in the vertices of the cells (Figs. 6 and 7). A more prolonged attack produces such concentration of points of attack in the vertices of the cells that they produce a distinct hole as shown in Figs. 8, 9, and 10.

The distribution of etch pits varies from one grain to another according to their crystalline orientation. We observe in some, starting from a distance of about 100 microns from the interface, a marked macromosaic substructure revealed by the alignment of etch pits. This substructure is closely linked with the cellular substructure (Hulme 1954; Takaka *et al.* 1955), because the etch pits usually pass through the vertices of the cells and keep a certain relation with their walls, though not following them closely (Figs. 8, 9, and 10). The size of the macromosaic substructure varies from that corresponding to one cell to that of various cells.

In accordance with the concepts set forth by Tiller (1958), the foregoing results may be interpreted as follows:

(1) In the zone of the interface and in the immediate surrounding zone, etch pits make it possible to detect the associations between impurities segregated in the borders of the cell substructure and dislocations formed by the mechanism of microsegregation of solute.

(2) The homogenization produced during the process of cooling liberates a great number of dislocations, especially those formed in the stages of growth of the interface and along the sides of the hexagonal substructure, which are eliminated by others of a contrary sign or reorder themselves following the mechanism of climbing suggested by Tiller (1958). Our results are in agreement with the recent work of Damiano and Tint in Zn (1961).

In strongly segregated zones (vertices of cells and zones of macrosegregation) neither homogenization nor migration of the grain boundary liberates anchored dislocations, for which reason attack with fluorated aqua regia will

reveal in almost every case the existence of etch pits in those zones of concentration.

Liberated dislocations, which have not been eliminated by others of a contrary sign, will tend to form minimum energy distributions (macromosaic substructure) and it seems logical to suppose that they have a connection with the existence of segregated zones, because the deformation energy of edge dislocations may be diminished by solute atoms, of a different size from that of the matrix, segregated in the dislocations (Cottrel 1948). The same mechanism may be applied in the case of alignments of dislocations forming the substructure of the macromosaic (Atwater and Chalmers 1957).

IV. CONCLUSIONS

Experiments with foundry polycrystals of 99.99% aluminum, in which solidification has been suspended by the sudden draining of the remaining liquid, furnish ample information about segregation substructures if we apply the method of epitaxial layers producing interference colors under oblique light.

By means of this technique it has been possible to determine the variation of segregation starting from the interface and proceeding towards the inner part of the grains. Detection, by means of thin epitaxial layers, of the persistence of segregation in the vertices of the cells and its disappearance in the borders of the hexagons seem to be conditioned by the orientations of the grains. Furthermore, it is possible to detect the existence of a "macrosegregation" substructure, which forms an enriched net of solute closely linked with the boundary of the grain. It corresponds in size to the original grain and within it there is a whole number of Smialowski cells.

The joint utilization of the method of etch pits and that of epitaxial layers makes it possible to establish: (1) distribution of the dislocation impurity associations starting from the interface and (2) their relation with segregated areas.

While in the interface the etch pits are concentrated along the borders of the cell; towards the inner grain they concentrate in macrosegregated areas, the vertices of the cells, and in alignments determining the substructure of the macromosaic. This substructure is closely linked with the vertices of the Smialowski cells and to a lesser degree with their walls. The interpretation of the distribution of corrosion figures is made in accordance with the theory on the formation of dislocations through microsegregations during solidification.

The effect of crystalline orientation is important on all the phenomena pointed out. For that reason we are at present engaged in experiments with single crystals of known orientation.

ACKNOWLEDGMENTS

Many thanks are extended to Dr. Gunther Schoeck and Professor Jorge A. Sabato for the interest they have shown in this work and for their very

valuable suggestions and to the National Council for Scientific and Technical Investigations for having subsidized this project.

REFERENCES

ATWATER, H. A. and CHALMERS, B. 1957. Can. J. Phys. **35**, 208.
BILONI, H. and DESTAILLATS, L. H. 1959. Métaux (Corrosion-Inds), **404**, 137.
BURTON, J. A., CABRERA, N., and FRANK, F. C. 1957. Phil. Trans. Roy. Soc. (London), A, **243**, 299.
COTTREL, A. H. 1948. Report Conf. on Strength of Solids, Phys. Soc. London, 24, 30.
DAMIANO, V. V. and TINT, G. S. 1961. Acta Met. **9**, 177.
FRANK, F. C. 1955. Deformation and flow of solids (IUTAM Colloquium, Madrid), p. 73.
FRIEDEL, J. 1956. Les dislocations (Gauthier-Villars, Paris).
HULME, X. 1954. Acta Met. **2**, 810.
JACQUET, P. 1956, 1957. Met. Revs. **1**, part 2.
LACOMBE, P. and MOUFLARD, M. 1953. Métaux (Corrosion-Inds), 340.
LACOMBE, P. and WYON, G. 1954. Lecture on defects in crystalline solids, Bristol, p. 347.
RUTTER, J. W. and CHALMERS, B. 1953. Can. J. Phys. **31**, 15.
SCHOECK, G. Private communication.
SCHOECK, G. and TILLER, W. A. 1960. Phil. Mag. **8**, 46.
TAKAKI, H., KOYAMA, X., and FUJIHIRA, H. 1955. Bull. Inst. Chem. Research, Kyoto Univ. **33**, 177.
TEGHTSOONIAN, E. and CHALMERS, B. 1951. Can. J. Phys. **29**, 370.
TILLER, W. A. 1958. J. Appl. Phys. **29** (4), 611.
TILLER, W. A. and RUTTER, J. W. 1956. Can. J. Phys. **34**, 96.
WALTON, D. and CHALMERS, B. 1959. Trans. AIME, **215** (3), 447.
WINEGARD, W. C. and CHALMERS, B. 1954. Trans. Am. Soc. Metals, **46**, 1214.
WYON, G., MARCHIN, X., and LACOMBE, P. 1959. Rev. mét. **56** (6).

NOTES

THE 3820-Å $^1A'' \leftarrow {}^1A'$ TRANSITION OF PROPYNAL

J. C. D. Brand,* J. H. Callomon,† and J. K. G. Watson*

The first electronic absorption band system in propynal $C_2H.CHO$ sets in at about 4200 Å (origin ca. 4143 Å)‡ and is extremely weak (ca. 1 m-atm of vapor). The second, rather stronger, system extends from about 3900 Å to 3000 Å, with electronic origin at 3821 Å. From their intensities, and by analogy with the spectra of other aldehydes (especially formaldehyde), it seems natural to identify the two systems with transitions respectively to the triplet and singlet states of the first excited electronic configuration. In formaldehyde, the effect of such excitation is to make the nuclear equilibrium configuration non-planar (Brand 1956; Robinson 1956) and it seemed of interest to see whether such behavior is retained in a substituted formaldehyde in which the substituent group would be expected to be able to conjugate (in the broadest sense) with the formyl carbonyl.

The 3820-Å system was observed at low dispersion by Howe and Goldstein (1958), who identified the probable electronic origin and tentatively assigned frequencies to some upper-state fundamentals. We have re-examined the spectrum comprehensively, at high dispersion, with a 6-meter Ebert grating spectrograph (King 1958) resolving 0.1 cm^{-1} in the second order, and have extended the measurements to include the $C_2D.CHO$ and $C_2H.CDO$ isotopic species. The molecule is very nearly a symmetric top, and the relatively simple rotational fine structure of the bands, both of K and J types, is well resolved (see Fig. 1). This allows the spectrum to be analyzed in rather greater detail than is perhaps usual in polyatomic molecules of this size, and as the complete analysis will take time, this note presents some of the preliminary results.

The spectral origin, 26163.1 cm^{-1} in the normal isotope, is marked by a strong perpendicular-type band shown by its rotational J-structure to be polarized along the inertial c-axis: thus the electronic transition is $A'' \leftarrow A'$ in notation appropriate to the ground state which is known to be planar from its microwave spectrum (Costain and Morton 1959). There is only one strong, long upper-state progression, also of type c bands, in a frequency identified with the C=O stretching vibration. Shorter type c progressions occur in other frequencies, and the difficulties so often encountered in distinguishing between bands originating from the zero vibrational level of the ground state and "hot" bands can here be avoided by the use of K-type combination differences obtained from ${}^RQ$ and ${}^PQ$ subband branches. In addition weak short progressions (of at most two members) are observed in frequencies in which

*Department of Chemistry, The University, Glasgow.
†Department of Chemistry, University College London, London, W.C.1.
‡This spectrum is being studied by G. W. King and D. Moule (personal communication)

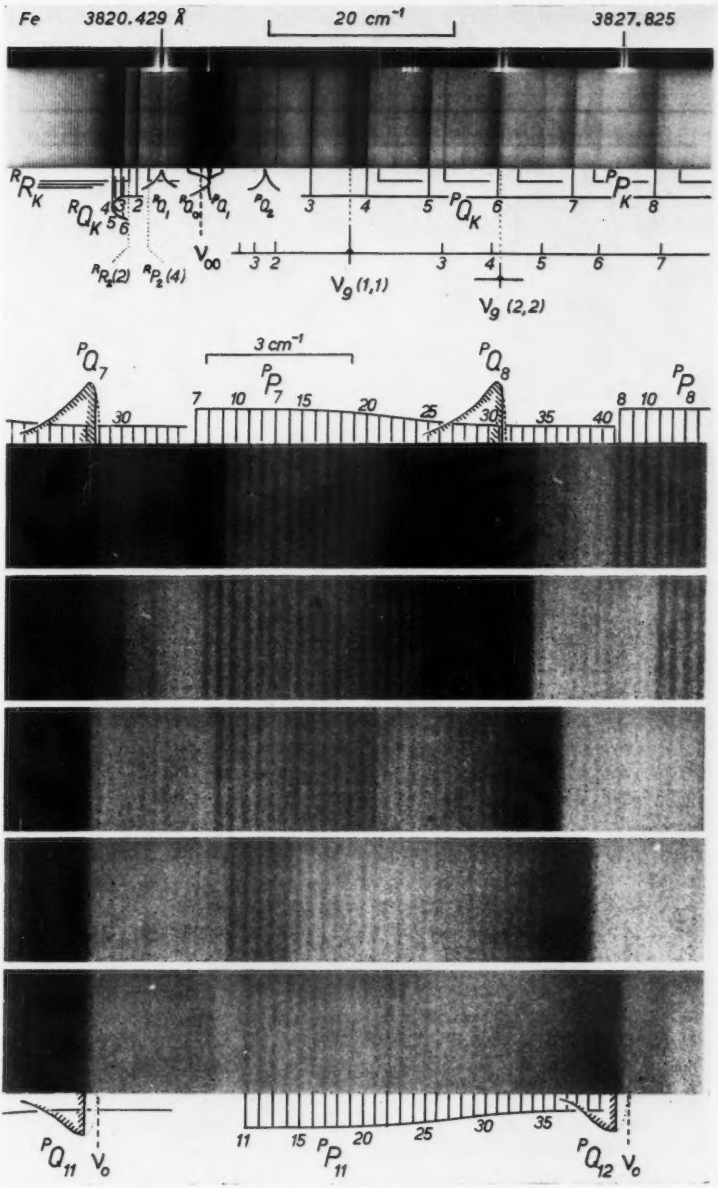


FIG. 1. The zero–zero absorption band of $C_2H.CHO$ near 3821 Å photographed in the second order of an Ebert grating spectrograph. Upper: appearance of the band around the band center. Lower: five successive P-type subbands ($\Delta K = -1$) placed side by side, showing details of the J-structure.

the first one-quantum members are type a – type b hybrid bands involving non-totally symmetric (out-of-plane) vibrations.

The analysis of the excited state vibration frequencies which have been identified in the normal molecule is summarized in the table, which also lists

TABLE I

Fundamental vibration frequencies of $C_2H.CHO$, in cm^{-1}

	Excited state	Ground state
Class a' (in-plane)		
ν_2 (formyl C—H stretch)	2952.9	2858.2[b]
ν_3 (C≡C stretch)	1945.7	2106[b]
ν_4 (C=O stretch)	1299.6[c]	1697
ν_5 (formyl C—H bend)	1119.5	1389[b]
ν_6 (C—C stretch)	957.1[d]	943.7
ν_8 (C—C=O bend)	506.7	614
ν_9 (C≡C—C bend)	189.2	205.5
Class a'' (out-of-plane)		
ν_{10} (formyl C—H deformation)	462.0	981.3
ν_{12} (skeletal deformation)	345.9	260.6

[a]Frequencies taken from measurements of central intensity maxima given to the nearest cm^{-1}.
[b]From the infrared; very weak or absent in the ultraviolet spectrum.
[c]Uncorrected for Fermi resonance with $(\nu_8+\nu_9)$ at 1311.5 cm^{-1}.
[d]Uncorrected for Fermi resonance with $2\nu_{10}$ at 943.6 cm^{-1}.

the analogous ground-state frequencies (values from the electronic spectrum agree with those from the infrared (Brand and Watson 1960; King and Moule 1961)). The vibrations which appear in the electronic spectrum are seen to be all those involving motion around the formyl carbon atom and include only one of the acetylenic vibrations. The most striking changes effect ν_4, which decreases in about the same proportion as in formaldehyde, and the two out-of-plane vibrations ν_{10} and ν_{12}. Any evidence of non-planarity of the excited state might therefore be sought in the latter. However, the progressions in these vibrations do not exceed at most two quanta, as might otherwise be expected on Franck–Condon grounds if a change of shape were involved. There is some doubt about the value of $2\nu'_{10}$, due to Fermi resonance, but its probable value of 943 cm^{-1} indicates some positive anharmonicity which is, however, still small compared with effects observed in the CH_2O spectrum. ν'_{12} can be observed up to three quanta in sequences, and behaves normally. Thus, from vibrational evidence, excited propynal is planar or very nearly planar in the equilibrium configuration; the height of any possible inversion barrier does not exceed a small fraction of the zero-point energy of ν'_{10}.

The rotational K-structure is strongly degraded to longer wavelengths and the J-structure very slightly degraded to shorter wavelengths. Analysis is straightforward, but has revealed a number of interesting perturbations. At least two types of Coriolis coupling have been well characterized. The one involves the ground-state frequencies ν_9 (a') and ν_{12} (a''), which were inaccessible in the infrared, and takes a form very similar to that identified there in ν_7 and ν_{11} (Brand and Watson 1960). $\zeta_{9,12}$ is close to 0.63. The other is found in $C_2H.CDO$. $2\nu'_4$ $(a', 2520\ cm^{-1})$ lies only a few cm^{-1} above another level, of

a'' symmetry whose rotational constant $(A—\bar{B})'$ differs such that its K-type energy levels rise more rapidly than, and overtake, those of $2\nu_4'$. Crossing occurs in the region of $K = 9$ where the levels mix equally, and the subbands of a single electronic band appear double. Rotational perturbations probably intermediate between these cases appear in other bands.

Finally, propynal shows an exceptionally clear-cut predissociation seen as an abrupt break in the subband structure of one perpendicular band. It occurs in transitions to the level ν_2' in $C_2H.CDO$; both ${}^PQ$ and ${}^PP$ branches in subbands up to $K' = 11$ are present normally, but ${}^PQ_{13}$ and ${}^PP_{13}$ branches are wholly missing. This resembles a similar case recently described by Clement and Ramsay (1961) in HNO. However, because of the smaller rotational constants, the predissociation limit can be narrowed down much more closely in the present case, for the energy difference between levels $K' = 11$ and $K' = 12$ is only 36 cm^{-1}. Another difference between HNO and propynal lies in the dependence of the predissociation on vibrational co-ordinates. In HNO, the limit is observed at the same energy no matter which vibrational state the molecule is in. In propynal the limit is clearly observed (as line-broadening) only in ν_2'. Bands involving transitions to other vibrational states at much higher energies retain their rotational fine structure in full. The predissociation limit gives an upper limit for the ground-state dissociation energy $D_0(D—CO.C_2H)$ of $28{,}595 \pm 17$ cm^{-1} (3.545 ev, 81.74 kcal/mole). From the extremely sharp onset of predissociation we conclude that the repulsive, predissociating potential curve must cut the ν_2' curve nearly horizontally, and that the upper limit to D_0 given above must be very close to the actual value.

We thank the Department of Scientific and Industrial Research for a maintenance grant (J. K. G. W.), and Dr. C. C. Costain for a sample of deuterated propynal.

Brand, J. C. D. 1956. J. Chem. Soc. 1956, 858.
Brand, J. C. D. and Watson, J. K. G. 1960. Trans. Faraday Soc. **56**, 1582.
Clement, M. J. Y. and Ramsay, D. A. 1961. Can. J. Phys. **39**, 205.
Costain, C. C. and Morton, J. R. 1959. J. Chem. Phys. **31**, 389.
Howe, J. A. and Goldstein, J. H. 1958. J. Am. Chem. Soc. **80**, 4846.
King, G. W. 1958. J. Sci. Instr. **35**, 11.
King, G. W. and Moule, D. 1961. Spectrochim. Acta, **17**, 286.
Robinson, G. W. 1956. Can. J. Phys. **34**, 699.

Received July 6, 1961.
Department of Chemistry,
The University, Glasgow,
and
Department of Chemistry,
University College London,
London, W.C.1.

THE IONIZATION POTENTIAL OF CH_2*

G. HERZBERG

A few years ago the ionization potential of CH_3 was determined from a Rydberg series in the vacuum ultraviolet. The result was 9.843 ev (Herzberg and Shoosmith 1956). When the first absorption bands of CH_2 were obtained in the flash photolysis of diazomethane near 1415 Å (Herzberg and Shoosmith 1959), it was suspected that they might represent the first member of a Rydberg series of CH_2, similar to the 1500-Å bands of CH_3. However, because of the shift to shorter wavelengths, it proved to be much more difficult to obtain higher members of the Rydberg series. By means of a grating, coated for high reflectivity in the 1200-Å region, it has now been possible to extend the CH_2 spectrum to shorter wavelengths and to observe higher members of the Rydberg series which make it possible to establish the ionization potential of CH_2.

As discussed elsewhere (Herzberg 1961) the 1415-Å group of CH_2 consists of a main band representing a $^3\Sigma_u^- - {}^3\Sigma_g^-$ transition and a number of line-like features at shorter wavelengths which, in all probability, represent two $^3\Pi - {}^3\Sigma$ transitions. On the new plates three further groups of the same type have been observed which together with the 1415-Å group form a clear Rydberg series. In Fig. 1, which shows the spectrum from 1324 to 1219 Å, the three groups are readily recognized. In addition, C I lines at 1329 and 1277 Å and CH bands at 1271 and 1234 Å are present, and, faintly, some C_2H_2 bands which are very prominent in the spectrum of the residue of the flash photolysis. As would be expected, the spacing of the individual subbands in each group decreases for the higher series members confirming that these subbands correspond to different electronic transitions with the same principal quantum number.

The first strong band in each group is listed in the second column of Table I. In the third column the wave numbers obtained from the formula

$$\nu = 83851 - [R/(n-0.12)^2]$$

are given. It is seen that this formula gives a very good representation of the observed bands confirming the Rydberg character of this series. Therefore the limit of the series 83851 cm^{-1} or 10.396 ev represents an ionization potential of the free CH_2 radical.

TABLE I

Rydberg series of CH_2

n	ν_{obs} (cm^{-1})	ν_{calc} (cm^{-1})
3	70634[a]	70621
4	76553	76562
5	79241	79243
6	80688	80677

[a]Band origin.

*Issued as N.R.C. No. 6502.

There remains the question of whether or not the ionization potential so obtained is the lowest ionization potential of CH_2. From the electron configuration of the upper state of the 1415-Å band it appears very likely that the main bands of the Rydberg series correspond to $2p\pi\ nd\pi\ {}^3\Sigma_u^-$ while the other features correspond to $2p\pi\ nd\sigma$ and $2p\pi\ nd\delta$. This interpretation is in agreement with the small size of the Rydberg correction. On this basis the state of the CH_2^+ ion to which the series converges is the $\sigma^2\sigma^2\sigma^2\pi\ {}^2\Pi$ state of the ion which is expected to be the ground state of the ion if it is linear. In a $^2\Pi$ state of a linear molecule we must expect a Renner–Teller splitting. If this splitting were large, as for example in NH_2 (Dressler and Ramsay 1959), it would mean that one component of the $^2\Pi$ state would have a potential minimum for a non-linear configuration, and this minimum would of course be below the minimum of the linear component. The limit of the Rydberg series must correspond to the linear component of $^2\Pi$ since only one strong band occurs in each member, that is, according to the Franck–Condon principle, the geometrical configuration of the molecule in the Rydberg states must be the same as in the ground state. Therefore, if there were a strong Renner–Teller interaction the true ionization potential of CH_2 would be somewhat smaller than the value given.

There are strong (but perhaps not entirely conclusive) reasons against the assumption that the ground state of CH_2^+ is non-linear and that therefore the true ionization potential of CH_2 is less than the value given.

(*a*) Transitions from the linear ground state to those components of the Rydberg states in which the molecule is bent should give rise to long progressions in the bending vibrations. No such progressions were found; but this may have been due to the low intensity of these bands.

(*b*) If CH_2^+ is bent in its ground state one would expect a Rydberg series of bent CH_2. As has been shown previously (Herzberg 1961) CH_2 is bent in the lowest singlet state ($^1A_1$) which can be observed in absorption at very short time intervals after the beginning of photolysis. Under these conditions no Rydberg series of singlet CH_2 was found. Again this could be due to low intensity but it would be difficult to see why a singlet Rydberg series should be much weaker than a triplet Rydberg series.

(*c*) Jordan and Longuet-Higgins (1961) on the basis of molecular orbital calculations similar to those which have given an excellent representation of the low excited states of CH_2, have concluded that CH_2^+ should be linear. This is also suggested by the fact that in the first excited singlet state of CH_2 ($a_1\,b_1\,{}^1B_1$) and in the first excited doublet state of NH_2 ($a_1\,b_1^2\,{}^2A_1$) the molecule is linear in spite of the fact that according to Walsh (1953) the $a_1\pi_u$ orbital tends to bend the molecule while $b_1\pi_u$ is neutral in this respect. In other words a single electron in the bending orbital (as in the ground state of CH_2^+) is not sufficient to make the molecule bent.

On the basis of these considerations it appears almost certain that the value given above represents the first ionization potential of CH_2. This value may be compared with the electron collision value of Langer, Hipple, and Stevenson (1954), who gave $11.7_5 \pm 0.1$ ev. The large discrepancy must be ascribed to the fact that in the electron collision experiments CH_2^+ is formed

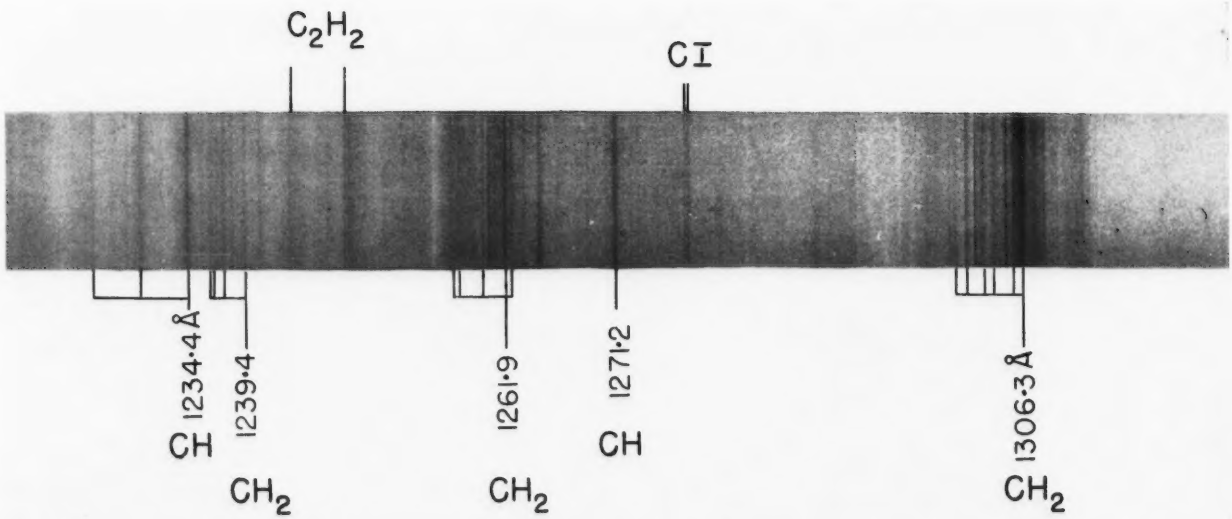


FIG. 1. Rydberg series of CH_2 in the vacuum ultraviolet. The first member at 1415 Å is not shown. Several CH bands as well as C I lines appear at the same time.

in an excited state with an excitation of 1.4 ev. There are several expected low-lying states of CH_2^+ from the configuration $\sigma\pi^2$. Of these the $^4\Sigma_u^-$ state is probably the lowest and might well lie at 1.4 ev.

If the new value for the ionization potential of CH_2 is combined with the appearance potential of CH_2^+ from CH_3 (viz. 15.3_0 ev) as obtained independently by Langer, Hipple, and Stevenson (1954) and Waldron (1954), one obtains immediately for the dissociation energy $D(CH_2\text{–}H)$ the value 4.9_0 ev (or 113 kcal). However, again one must consider the possibility that CH_2^+ is formed in an excited state and therefore the $D(CH_2\text{–}H)$ value just given can only be considered as an upper limit for the true value. From it and the heat of formation of CH_4 together with $D(CH_3\text{–}H)$ and $D(CH)$ one can immediately derive a lower limit for $D(CH\text{–}H)$, namely 98 kcal.

I am greatly indebted to Mr. J. Shoosmith for obtaining the plates on which this discussion is based, and to Drs. F. P. Lossing and A. E. Douglas for useful criticisms.

Dressler, K. and Ramsay, D. A. 1959. Trans. Roy. Soc. **251**, 553.
Herzberg, G. 1961. Proc. Roy. Soc. A, **262**, 291.
Herzberg, G. and Shoosmith, J. 1956. Can. J. Phys. **34**, 523.
——— 1959. Nature, **183**, 1801.
Jordan, P. C. H. and Longuet-Higgins, H. C. To be published.
Langer, A., Hipple, J. A., and Stevenson, D. P. 1954. J. Chem. Phys. **22**, 1836.
Waldron, J. D. 1954. Trans. Faraday Soc. **50**, 102.
Walsh, A. D. 1953. J. Chem. Soc. 2260.

Received July 14, 1961.
Division of Pure Physics,
National Research Council,
Ottawa, Canada.

Under this heading brief reports of important discoveries in physics may be published. These reports should not exceed 600 words and, for any issue, should be submitted not later than six weeks previous to the first day of the month of issue. No proof will be sent to the authors.

Magnetization Nutations in Maser Oscillation Pulses

The amplitude-modulated structure sometimes observed in oscillation pulses from two-level solid-state masers has been investigated theoretically in several reports. In two of these (Kemp 1959; Yariv 1960), it has been fairly well established that this modulation can be attributed to a large-amplitude nutation of the magnetization vector M_0. If this is the case, it would be reasonable to expect the r-f. paramagnetic susceptibility to remain a pure imaginary quantity (at resonance) for the duration of the pulse. Thus, in the case of maser oscillation in a resonant cavity, the susceptibility component χ'' along the imaginary axis would either be negative or positive, depending on whether the specimen was emitting power ($dM_z/dt > 0$) to, or receiving power ($dM_z/dt < 0$) from, the cavity. M_z is the longitudinal component of volume magnetization, equal to $(+M_0)$ in the steady state.

This situation is self-evident in the model proposed by Kemp (1959). It is not evident, however, in Yariv's (1960) relationships. This paper is concerned with an examination of the latter relationships for disclosure of this feature.

Yariv's (1960) equations are, in Gaussian units,

$$aM_r + b\frac{dM_i}{dt} + Q\frac{dH_i}{dt} + \omega H_i = 0, \tag{1}$$

$$-aM_i + b\frac{dM_r}{dt} + Q\frac{dH_r}{dt} + \omega H_r = 0, \tag{2}$$

$$\frac{dM_i}{dt} + |\gamma|M_z H_r + CM_i = 0, \tag{3}$$

$$\frac{dM_r}{dt} - |\gamma|M_z H_i + CM_r = 0, \tag{4}$$

$$\frac{dM_z}{dt} + |\gamma|M_r H_i - |\gamma|M_i H_r = 0, \tag{5}$$

where $a = 4\omega\eta Q$, $b = a/\omega$, $c = T_2^{-1}$, and M_r, M_i and H_r, H_i are the real and imaginary components of the transverse magnetization and r-f. magnetic field, respectively. $\gamma = -|\gamma|$ is the electron gyromagnetic ratio, η is the sample filling factor, and Q is the loaded "Q" of the cavity. $f = \omega/2\pi$ is the cavity resonance and paramagnetic transitional frequency in this case. T_2 is the spin–spin relaxation time for the paramagnetic specimen. The spin–lattice relaxation time T_1 is tacitly assumed to be infinite.

The chosen boundary conditions were $M_z(0) = -M_0$, $M_r(0) = 0$, $M_i(0) = 0$, $H_r(0) = 0.1\ M_0$, $H_i(0) = 0$. M_0 could be evaluated from a knowledge of the F-center concentration, splitting-field magnitude, and operational temperature. Note that a perturbation term (such as H_r) is necessary if the system is to depart from a state of unstable equilibrium. Selected values were:

$$\begin{aligned} M_0 &= 5\times10^{-4}\ \text{gauss},\\ H_r(0) &= 5\times10^{-5}\ \text{gauss},\\ Q &= 10^4,\\ \eta &= 0.1,\\ f &= 10^{10}\ \text{c.p.s.},\\ T_2^{-1} &= 0, \end{aligned}$$

where the last term corresponds to a negligible spin–spin interaction.

The five time-dependent variables were evaluated by the author with the aid of a Bendix G-15 digital computer. The Runge–Kutta–Gill (1951) method was employed. The variables were evaluated at intervals of 2.0×10^{-9} seconds over a range of 1.0×10^{-6} seconds. Values for $|M(t)|$ and $|H(t)|$ were then obtained (where $M^2 = M_r^2 + M_i^2$, $H^2 = H_r^2 + H_i^2$) together with the phase angle between M and H vectors. It was found that a 90° phase shift was indeed preserved between the two variables M and H, with M leading for $dM_z/dt > 0$ and M lagging for $dM_z/dt < 0$. The time-dependent behavior of M_z and H^2 was similar to that shown by Yariv, with the exception that only one zero-crossing of the M_z component was found for the parameters employed. In addition, the peaks of M_z were displaced very slightly

from those of H^2 ($\Delta t \sim 8\times10^{-9}$ seconds). The spin trajectories in a rotating co-ordinate frame would thus be in agreement with those discussed by Kemp (1960) for a weak spin–spin interaction. In fact, when Kemp's equations were also programmed on the computer, the derived time-dependent magnetization and r-f. magnetic field patterns were of similar shape and duration to those shown by Yariv, for similar boundary conditions and circuit parameters.

The author wishes to thank Mr. I. Firth, Department of Natural Philosophy, St. Andrews University, Scotland, for many valuable suggestions on this topic.

GILL, X. 1951. Proc. Cambridge Phil. Soc. **47**, 96.
KEMP, J. C. 1959. J. Appl. Phys. **30**, 1451.
YARIV, A. 1960. J. Appl. Phys. **31**, 740.

RECEIVED MAY 15, 1961.
DEPARTMENT OF ELECTRICAL ENGINEERING,
MCMASTER UNIVERSITY,
HAMILTON, ONTARIO.

C. K. CAMPBELL

An Observation of Noctilucent Clouds in Western Canada

Only one authentic report of noctilucent clouds over Canada is known to the writers (Vestine 1934). This may be compared with the 10 to 30 reported occurrences each year in the European part of the U.S.S.R. (Grishin 1957) and the systematic studies on the luminous clouds made in the Scandinavian Peninsula (Witt 1957). An examination of the 16-mm, all-sky camera records of aurora over West-central Canada, taken during the I.G.Y., failed to disclose cloud structures of this type. Since there is much distortion in the portions of the photographs close to the horizon, it is difficult to distinguish between a luminous cloud and faint aurora; and this in itself is not conclusive proof that the clouds were not present on many occasions.

Possible reasons for the lack of observations of noctilucent clouds over Canada were discussed at a seminar early in July. Some of the reasons advanced were:

(*a*) failure to make observations at times and in directions most favorable to their detection;

(*b*) confusion between the luminous clouds and weak aurora;

(*c*) significant differences in the upper atmospheric circulations between Northern Canada and Europe.

The clouds are observed most frequently when the sun is from 6 to 18° below the horizon, at altitudes below 6°, and at latitudes north of 50° during the summer months. The low altitudes for maximum visibility, and in directions toward the sun and the zone of maximum auroral occurrence for Canada, make it easy for an inexperienced observer to designate all faint luminous structures as auroral in origin. A notable difference in the regional occurrence of the clouds, if it actually exists, might help to explain their origin.

Probably as a result of these discussions, a display of the noctilucent clouds was observed on the morning of July 24. A number of photographs were taken. The clouds had the tenuous and shiny white appearance described by European observers. The solar depression during the period of observation varied between 6 and 11°. The major portions of the visible clouds were below an angular altitude of 6° and in azimuths toward the sun.

This observation together with the special circumstances which induced a more intensive search for noctilucent clouds suggest that they appear with much greater frequency over Northern Canada than the number of reported occurrences suggest. While it is now too late in the year for systematic observational programs to be started, it is suggested that such programs should be initiated in future years.

GRISHIN. 1957. Izvest. Akad. Nauk S.S.S.R., Moscow.
VESTINE. 1934. J. Roy. Astron. Soc. **28**, 249.
WITT. 1957. Tellus, **9**, 341.

RECEIVED AUGUST 9, 1961.
INSTITUTE OF UPPER ATMOSPHERIC PHYSICS,
PHYSICS DEPARTMENT,
UNIVERSITY OF SASKATCHEWAN,
SASKATOON, SASKATCHEWAN.

K. D. BAKER
B. W. CURRIE

NOTES TO CONTRIBUTORS

Canadian Journal of Physics

MANUSCRIPTS

General.—Manuscripts, in English or French, should be typewritten, double spaced, on paper 8½×11 in. **The original and one copy are to be submitted.** Tables and captions for the figures should be placed at the end of the manuscript. Every sheet of the manuscript should be numbered. Style, arrangement, spelling, and abbreviations should conform to the usage of recent numbers of this journal. Greek letters or unusual signs should be written plainly or explained by marginal notes. Characters to be set in boldface type should be indicated by a wavy line below each character. Superscripts and subscripts must be legible and carefully placed. Manuscripts and illustrations should be carefully checked before they are submitted. Authors will be charged for unnecessary deviations from the usual format and for changes made in the proof that are considered excessive or unnecessary.

Abstract.—An abstract of not more than about 200 words, indicating the scope of the work and the principal findings, is required, except in Notes.

References.—References should be listed **alphabetically by authors' names**, unnumbered, and typed after the text. The form of the citations should be that used in current issues of this journal; in references to papers in periodicals, titles should not be given and only initial page numbers are required. The names of periodicals should be abbreviated in the form given in the most recent *List of Periodicals Abstracted by Chemical Abstracts.* All citations should be checked with the original articles and each one referred to in the text by the authors' names and the year.

Tables.—Tables should be numbered in roman numerals and each table referred to in the text. Titles should always be given but should be brief; column headings should be brief and descriptive matter in the tables confined to a minimum. Vertical rules should not be used. Numerous small tables should be avoided.

ILLUSTRATIONS

General.—All figures (including each figure of the plates) should be numbered consecutively from 1 up, in arabic numerals, and each figure referred to in the text. The author's name, title of the paper, and figure number should be written in the lower left corner of the sheets on which the illustrations appear. Captions should not be written on the illustrations.

Line drawings.—Drawings should be carefully made with India ink on white drawing paper, blue tracing linen, or co-ordinate paper ruled in blue only; any co-ordinate lines that are to appear in the reproduction should be ruled in black ink. Paper ruled in green, yellow, or red should not be used. All lines must be of sufficient thickness to reproduce well. Decimal points, periods, and stippled dots must be solid black circles large enough to be reduced if necessary. Letters and numerals should be neatly made, preferably with a stencil (**do NOT use typewriting**) and be of such size that the smallest lettering will be not less than 1 mm high when the figure is reduced to a suitable size. Many drawings are made too large; originals should not be more than 2 or 3 times the size of the desired reproduction. Whenever possible two or more drawings should be grouped to reduce the number of cuts required. In such groups of drawings, or in large drawings, full use of the space available should be made; the ratio of height to width should conform to that of a journal page (4¾×7½ in.), but allowance must be made for the captions. **The original drawings and one set of clear copies (e.g. small photographs) are to be submitted.**

Photographs.—Prints should be made on glossy paper, with strong contrasts. They should be trimmed so that essential features only are shown and mounted carefully, with rubber cement, on white cardboard, with no space between those arranged in groups. In mounting, full use of the space available should be made. **Photographs are to be submitted in duplicate;** if they are to be reproduced in groups one set should be mounted, the duplicate set unmounted.

REPRINTS

A total of 100 reprints of each paper, without covers, are supplied free. Additional reprints, with or without covers, may be purchased at the time of publication.

Charges for reprints are based on the number of printed pages, which may be calculated approximately by multiplying by 0.6 the number of manuscript pages (double-spaced typewritten sheets, 8½×11 in.) and including the space occupied by illustrations. Prices and instructions for ordering reprints are sent out with the galley proof.

Contents

Printed in Canada *by* University of Toronto Press